सत्यमेव जयते

INDIAN AGRICULTURAL
RESEARCH INSTITUTE, NEW DELHI

14704

I.A.R I.6.
GIP NLK—H 3 I.A.R.I.—10-5-55—15,000

THE
PROCEEDINGS
OF
THE PHYSICAL SOCIETY

FROM JANUARY 1942 TO NOVEMBER 1942

VOLUME 54

Published by
THE PHYSICAL SOCIETY
1 Lowther Gardens, Exhibition Road,
London, S.W. 7

Printed by
TAYLOR AND FRANCIS, LTD.,
RED LION COURT, FLEET STREET, LONDON, E.C.4

14704

OFFICERS AND COUNCIL, 1942–43

President:

Sir CHARLES G. DARWIN, K.B.E., M.C., M.A., Sc.D., F.R.S.

Vice-Presidents who have filled the Office of President:

Sir CHARLES VERNON BOYS, LL.D., F.R.S.
C. H. LEES, D.Sc., F.R.S.
ALEXANDER RUSSELL, M.A., D.Sc., F.R.S.
Sir FRANK E. SMITH, K.C.B., G.B.E., D.Sc., F.R.S.
Sir OWEN W. RICHARDSON, M.A., D.Sc., F.R.S.
W. H. ECCLES, D.Sc., M.I.E.E., F.R.S.
Sir ARTHUR S. EDDINGTON, O.M., M.A., D.Sc., F.R.S.
A. O. RANKINE, O.B.E., D.Sc., F.R.S.
The Right Hon. Lord RAYLEIGH, M.A., Sc.D., LL.D., F.R.S.
T. SMITH, M.A., F.R.S.
ALLAN FERGUSON, M.A., D.Sc.

Vice-Presidents:

G. I. FINCH, M.B.E., D.Sc., F.R.S.
J. H. BRINKWORTH, D.Sc.
C. D. ELLIS, B.A., Ph.D., F.R.S.
H. T. FLINT, D.Sc., Ph.D.

Hon. Secretaries:

W. JEVONS, D.Sc., Ph.D. (*Business*)
J. H. AWBERY, B.A., B.Sc. (*Papers*)

Hon. Foreign Secretary:

Sir OWEN W. RICHARDSON, M.A., D.Sc., F.R.S.

Hon. Treasurer:

C. C. PATERSON, O.B.E., D.Sc., M.I.E.E., F.R.S.

Hon. Librarian:

L. C. MARTIN, D.Sc.

Ordinary Members of Council:

N. F. MOTT, M.A., F.R.S.
J. NICOL, B.A., B.Sc.
D. OWEN, B.A., D.Sc.
Sir EDWARD V. APPLETON K.C.B., M.A., D.Sc., LL.D., F.R.S.
Miss M. D. WALLER, Ph.D.
E. N. da C. ANDRADE, D.Sc., Ph.D., F.R.S.
H. SHAW, D.Sc.
E. R. DAVIES, B.Sc.
W. B. MANN, Ph.D.
A. J. PHILPOT, O.B.E., M.A., B.Sc.
H. C. WEBSTER, M.Sc., Ph.D.
W. D. WRIGHT, D.Sc.

Editors:

Proceedings:
J. H. AWBERY, B.A., B.Sc.

Reports on Progress in Physics:
W. B. MANN, Ph.D.

COLOUR GROUP

Chairman:
W. D. WRIGHT, D.Sc.

Hon. Secretary:
H. D. MURRAY, M.A., F.I.C.

OPTICAL GROUP

Chairman:
A. O. RANKINE, O.B.E., D.Sc., F.R.S.

Hon. Secretary:
L. C. MARTIN, D.Sc.

CONTENTS

Part 1. 1 *January* 1942

PAGE
J. G. GRUNDY. Problems of colour mixing in the dyeing industry 1
ROBERT SCHNURMANN and ERIC WARLOW-DAVIES. The electrostatic components of the force of sliding friction 14
H. R. NETTLETON and E. G. BALLS. The absolute measurement of electrical resistance by a method using the average electromotive force of a commutating generator 27
G. D. YARNOLD. Note on electric and magnetic dimensions 46
C. A. COULSON. Diamagnetic susceptibility of methane 51
V. D. HOPPER. Thermostats employing external surface control 55
S. WHITEHEAD. An approximate solution for the distribution of temperature or potential with cylindrical isothermal or equipotential surfaces 63
J. T. MACGREGOR-MORRIS and A. G. STAINSBY. A rotating differential photoelectric photometer for precision work. (Discussion.) 66
Reviews of books 70
Recent reports and catalogues 80

Part 2. 1 *March* 1942

J. G. HOLMES. The nature and measurement of whiteness 81
V. G. W. HARRISON. The measurement of near-whites in the paper industry . 86
C. G. HEYS HALLETT. The whiteness of cinema screens 98
D. E. M. GARFITT. The measurement of impulse voltages by means of small sphere gaps 109
W. E. BENHAM. The nature of temperature 121
H. H. MACEY. A problem in two-dimensional flow 128
W. H. J. FUCHS and R. KOMPFNER. On space-charge effects in velocity-modulated electron beams 135
C. W. HANSEL. Fundamental laws and definitions in physics : I—Ohm's law . 151
C. W. HANSEL. Fundamental laws and definitions in physics : II—Specific heat and Newton's law of cooling 159
C. W. HANSEL. Fundamental laws and definitions in physics : III—Mass . . 164
Reviews of books 205

Part 3. 1 *May* 1942

William David Coolidge : Eighteenth Duddell Medallist 209
DAVIS S. EVANS and K. MENDELSSOHN. The physics of blood-pressure measurement 211
F. H. G. PITT. Colour blindness and its importance in relation to industry . . 219
J. D. CRAGGS. Canal-ray positive-ion sources 245
R. E. SIDAY. The determination of the optical properties of thick magnetic lenses, and the application of these lenses to beta-ray spectrometry 266
G. F. C. SEARLE. The determination, by smoke manometer, of the ratio of density of town gas to that of air at the same temperature and pressure . . . 278
B. K. JOHNSON. Demonstration. Photomicrographs illustrating the paper " An achromatic reflection microscope for use with visible or ultra-violet light " . 288
Discussion on papers by W. E. BENHAM and G. D. YARNOLD entitled " On the nature of temperature " and " Notes on electric and magnetic dimensions ", respectively 289

PAGE

Obituary notices :

JULIUS RHEINBERG 297

MARK BENJAMIN 297

TOM HARRIS 298

Reviews of books 300

Recent reports and catalogues 302

Part 4. 1 *July* 1942

W. D. WRIGHT. Research on colour physics at South Kensington, 1877–1942 . 303

KATHLEEN LONSDALE. X-ray study of crystal dynamics : An historical and critical survey of experiment and theory 314

(The late) Sir WILLIAM BRAGG. The secondary x-ray spectrum of sylvine : Theory and experiment 354

MAX BORN. Lattice dynamics and x-ray scattering 362

M. BLACKMAN. Note on Raman's theory of the specific heat of solids . . . 377

A. CHARLESBY. Note on the comparison of the subsidiary maxima in electron diffraction and x-ray diffraction patterns from single crystals 379

Recent reports and catalogues 387

Reviews of books 388

Part 5. 1 *September* 1942

W. M. HAMPTON. Some problems relating to optical glass 391

W. V. MAYNEORD. The measurement of radiation for medical purposes . . 405

H. FRÖHLICH. Dielectric loss in paraffin-wax solutions 422

E. W. H. SELWYN. The dimensions of physical quantities 428

F. T. FARMER. An electrometer for measurement of voltage on small ionization chambers 435

J. D. CRAGGS. A high-voltage apparatus for atomic disintegration experiments . 439

R. C. EVANS and H. S. PEISER. A machine for the computation of structure factors 457

Reviews of books 463

Part 6. 1 *November* 1942

A. CHARLESBY. Influence of barrier dimensions on the vibrational spectrum of ethane 471

A. J. C. WILSON. The thermal expansion of aluminium : Further experiments . 487

Sir ARTHUR EDDINGTON. The theoretical values of the physical constants . . 491

W. E. DUNCANSON. Note on some recent papers on physical quantities and their dimensions 504

Discussion on papers by C. W. HANSEL entitled " Fundamental laws and definitions in physics " 509

Presentation of Duddell Medal 527

Obituary notices :

Sir WILLIAM BRAGG 531

WILLIAM ALFRED BENTON. 535

FRANCIS JAMES SELBY 536

BERNARD MYNOTT NEVILLE 537

Reviews of books 538

Recent reports and catalogues 541

Index to volume 54. 542

Index to reviews of books, volume 54 546

PROCEEDINGS AT THE MEETINGS OF THE PHYSICAL SOCIETY

SESSION 1941–42

24 *September* 1941

The fourth meeting of THE COLOUR GROUP. At the Imperial College, London, S.W. 7, Dr. W. D. Wright being in the Chair.

The following three papers on Whiteness were read and were followed by an informal discussion :

(i) " The nature and measurement of whiteness ", by J. G. Holmes, B.Sc.

(ii) " The measurement of near-whites in the paper industry ", by V. G. W. Harrison, Ph.D.

(iii) " The apparent whiteness of cinema screens ", by C. G. Heys Hallett, M.A.

10 *October* 1941

At the Imperial College, London, S.W. 7, the Acting President, Professor Allan Ferguson, being in the Chair.

The following were elected to Fellowship, the first five being transferred from Student Membership : Percy George Forsyth, Paul Howard-Flanders, Dennis Albert Nicholls, Lawrence William Walker, John Bernard Warren, Albert Edward Bishop, Patrick Moran, Stanley Austen Stigant, Hardwicke Slingsby Tasker, Walter Harold Ward, George Stuart James White.

It was announced that Council had elected the following to Student Membership : Margaret Elisabeth Jane Carr, Alan Stephen Edmondson, Joseph William Fox, George Francis Hodsman and Denys Haigh Wilkinson.

A lecture on " Recent work on the use of photo-electric rectifier-type cells in photometry " was delivered by Professor J. T. MacGregor-Morris, D.Sc., M.I.E.E., and was followed by an informal discussion.

10 *December* 1941

The fifth meeting of THE COLOUR GROUP. At the Lighting Service Bureau of the Electric Lamp Manufacturers' Association, London, W.C. 2, Dr. W. D. Wright being in the Chair.

A lecture on " Colour blindness and its importance in relation to industry " was delivered by F. G. H. Pitt, Ph.D., and was followed by an informal discussion.

The meeting was preceded by an Extraordinary General Meeting of the Group for the consideration and adoption of an amendment of a clause in the Constitution enlarging the Committee of the Group.

16 *January* 1942

At the Imperial College, London, S.W. 7, Sir William Bragg, Past President, being in the Chair.

The following were elected to Fellowship : George Keith Thurburn Conn, Stanley Walter Cousins (transferred from Student Membership), Cornelius Frank Dietrich, Trevor Anthony Eames, Eric Leopold Irwin Heinzelmann, R. G. Horner, Ronald Huby, Alan Orville Hunter, Basil Jacob, Sybil Mathews, Oliver Moriogh Moriarty, Robert Kenworthy Schofield, Paul Vigoureux, Felix Bernard White.

It was announced that Council had elected the following to Student Membership : Harold Raynor Allan, Edward Raymond Andrew, Kenneth Robert Atkins, Stanley Reginald Badcoe, Norman Rushton Bailey, John Robert Barker, Roderick K. Barnes, Clifford Henry James Beaven, Arthur Reginald Bevan, Margaret Joy Boyes-Watson, Roy Drakeley Brace, F. J. Bradshaw, A. P. Brain, David Brechner, R. I. B. Cooper, D. W. Davies, David John Day, Philip W. Day, Jack Dunkley, Frederick Henry East, David William Elson, Reginald Francis Farr, Geoffrey Charles Frederick Fisher, S. G. F. Frank, Anthony Philip French, R. M. Goody, J. Hawkins, Reginald Dennis Hayward, R. Heastie, Kenneth William Hillier, Francis Raymond Holt, Denys Stanley Hopper, Howard Arthur Hughes, Robert W. G. Hunt, William Alan Jennings, Ronald Frederick Johnston, Robert Henry Kay, John Edward Anthony Keen, Derek LaWindell Watts King, Peter Reginald Layton, George Matthew Leak, Ian Alexander Darroch Lewis, Keith B. Logie, Samuel Thomas Lunt, George Albert Mann, Wilfred John Mellors, James Woodham Menter, Robert Henry Merson, Thomas Alan Minns, Eric Norbury Mutch, Audrey Mary Brasnett Parker, John Robert Pollard, Ernest Henry Putley, Alan Purcell Roberts, Henry Sheppard, Norman Peter Staines, James Richard Stansfield, Kenneth H. Stewart, Alan Stott, Bernard Sugarman, Charles Alfred Taylor, David West, Douglas Francis White, Rudolf Arnold Wiersma, John Charles Willmott, Laurence William Stevens Wilson, Robert Lewis Woolley, Anthony Frank Woolner.

" A lecture on " X-ray study of crystal dynamics " was given by Mrs. Kathleen Lonsdale, D.Sc., and was followed by an informal discussion.

---

21 *January* 1942

A joint meeting with the Cambridge Philosophical Society at the Cavendish Laboratory, Cambridge, the Acting President, Professor Allan Ferguson, being in the Chair.

Dr. Kathleen Lonsdale repeated the lecture on " X-ray study of crystal dynamics " (in a modified form). The lecture was followed by an informal discussion.

---

6 *February* 1942

At the Imperial College, London, S.W. 7, the Acting President, Professor Allan Ferguson, being in the Chair.

The following were elected to Fellowship, the first seventeen being transferred from Student Membership : Maurice Darley Armitage, E. G. S. Blower, Philip Charles Bowes, Helen Sarah Griffin, Philip Halliday, A. Hammond, Robert James Hercock, Arthur Cambridge Merrington, David Colver Nutting, William Thomas Shipston Pearson, Willaim Roy Piggott, James Reekie, Raymond Frederick Saxe, Anthony Charles Shearman, Andrew Stratton, John Carruthers Weston, K. B. S. Willder, Henry Shull Arms, Robert Sidney Blackledge, Charles Edward Rhodes Bruce, Edwin Oswald Cook, Cyril Bertie Daish, Lionel James Davies, Arthur E. Evans, O. R. Frisch, William Reginald Stephen Garton, Kurt Martin Guggenheimer, Henry George William Harding, Vernon George Wentworth Harrison, Nicholas Kurti, John Mark Anthony Lenihan, Robert Schnurmann, Poul Arne Scott-Iversen, Alfred Morris Thomas, G. A. Veszi, William Henry Walton, William Harford Willott, Arthur James Cochran Wilson, W. Rees Wright.

The following papers were read and discussed :

" Electrical and magnetic dimensions ", by G. D. Yarnold, M.A., D.Phil.

" The nature of temperature ", by W. E. Benham, B.Sc.

---

11 *February* 1942

The sixth meeting of THE COLOUR GROUP. At the Imperial College, London, S.W. 7, Dr. W. D. Wright being in the Chair.

Dr. W. D. Wright, as Chairman of the Colour Group, delivered his address entitled " Research on Colour Physics at South Kensington, 1877–1942 ".

27 *February* 1942

At the Imperial College, London, S.W. 7, the President, Dr. C. G. Darwin, being in the Chair.

The following were elected to Fellowship, the last-named being transferred from Student Membership : Samuel John Martin, Edward Frank Michaelson and Ronald Charles Pankhurst.

A lecture on " Measurement of radiation for medical purposes " was delivered by Professor W. V. Mayneord, D.Sc., and was followed by an informal discussion.

6 *March* 1942

The first meeting of THE OPTICAL GROUP. At the Science Museum and the Imperial College, London, S.W. 7, Dr. A. O. Rankine being in the Chair.

A Draft Constitution of the Optical Group as drawn up by the Committee of the Group was discussed and adopted. The Chairman, Honorary Secretary, and Committee of the Group for 1942–43 were elected.

A lecture on " Problems relating to optical glass " was delivered by W. M. Hampton, Ph.D., and was followed by an informal discussion.

The following papers were read and discussed :

" A multi-purpose collimator ", by R. J. Bracey, and

" Thermal effects on the performance of lens systems ", by J. W. Perry.

20 *March* 1942

At the Science Museum, London, S.W. 7, the President, Dr. C. G. Darwin, being in the Chair.

The following were elected to Fellowship : Fred Bovey, John Crank, Bertram John Frederick Dorrington, Walter Stevenson Joyce, John Smith Macpherson, Alex Renfrew Miller (transferred from Student Membership), George Mole, Dorothy L. Tilleard.

It was announced that Council had elected the following to Student Membership : Herbert Cairns Bolton, Kenneth John Carpenter, Norman Cusack, Grahame John Dickins, Gareth Wyn Evans, Douglas Hugh Everett, John Gadsby, Geoffrey Blakeley Greenough, M. H. James Hawkins, Frederick G. Hellyer, James Morden Hough, Fathi As'ad Qaddura, Alan Gerald Senior, J. R. Ennis Smith, Alexander R. Stokes, Ronald Walter Tiffen, Nigel Graham Trott, Theodore James Tulley, Anthony H. Willbourn, Donald Williams.

The twenty-sixth Guthrie Lecture was delivered by Sir Edward Appleton, K.C.B., M.A., D.Sc., LL.D., F.R.S., who took as his subject " Ionospheric influences on geomagnetism ".

27 *March* 1942

The seventh meeting of THE COLOUR GROUP. At the Midland Hotel, Manchester, Professor Allan Ferguson, Past President, being in the Chair.

A paper on " The physical significance of the dyers' system of colour matching ", by G. S. J. White, B.A., T. Vickerstaff, Ph.D., and E. Waters, B.Sc., was read and discussed.

9 *April* 1942

A joint meeting with the Science Masters' Association at Rugby School, Professor Allan Ferguson, Past President, being in the Chair.

A discussion was held on certain aspects of the teaching of Physics in schools, the opening papers being :

" Fundamental laws and definitions in physics : I—Ohm's law ; II—Specific heat and Newton's law of cooling ; III—Mass ", by C. W. Hansel, B.Sc.

" The need for a permanent standardization and guidance committee for elementary physics ", by P. Woodland, M.A.

15 *May* 1942

The second meeting of THE OPTICAL GROUP. At the Imperial College, London, S.W. 7, Dr. A. O. Rankine being in the Chair.

A demonstration of apparatus for micro-film processes was given by B. K. Johnson.

The following papers were read and discussed:

" A method of focusing photographic lenses ", by E. W. H. Selwyn, B.Sc.

" A spectrohelioscope ", by the Rev. W. Rees Wright, M.Sc., F.R.A.S.

A lecture on " The principles of the use of non-reflecting films in optical instruments " was delivered by K. M. Greenland, Ph.D.

---

20 *May* 1942

A joint meeting with the Royal Meteorological Society at the Imperial College, London, S.W. 7, the Chair being taken first by Professor G. I. Finch, Vice-President, and afterwards by Professor D. Brunt, President of the Royal Meteorological Society.

The following were elected to Fellowship: Dorothy Alice Baker, George Henry Giles, James Arthur Gordon, Cyril Harold Marsh, George Robert Noakes, Leonard Rotherham.

It was announced that the Council had elected the following to Student Membership: Harry E. Airey, John Geoffrey Dawes, Arthur Dennis Kent, Alan Denis McQuillan, Arthur James Shiell, Harry Shorland.

A discussion was held on " Emission and absorption of radiation in the atmosphere ", the opening paper being by Dr. T. G. Cowling and further contributions by Dr. G. M. B. Dobson, Professor H. Dingle, Dr. R. W. B. Pearse, Dr. A. G. Gaydon, Dr. L. Kellner, Dr. G. B. B. M. Sutherland, Professor F. A. Paneth, Professor P. A. Sheppard, Mr. E. Gold and Professor D. Brunt.

---

27 *May* 1942

The eighth meeting of THE COLOUR GROUP. At the Royal Photographic Society, London, S.W. 7, Dr. W. D. Wright being in the Chair.

Dr. R. A. Houstoun's colour-vision apparatus was demonstrated by J. W. Perry.

A paper on the " Relative merits of spectrophotometry and colorimetry " was read by Miss Dorothy L. Tilleard, B.Sc., and was followed by an informal discussion.

---

5 *June* 1942

At the Lighting Service Bureau of the Electric Lamp Manufacturers' Association, London, W.C. 2, Dr. A. O. Rankine, Chairman of the Optical Group and Past President of the Society, being in the Chair.

The following were elected to Fellowship: John Frederick William Bell, Edmund John Bowen, Hugh Mills Bunbury, Donald Anderson Crooks, George Henry Morris Farley, Kenneth Joseph Habell, William Francis Holford, Stewart Paterson (transferred from Student Membership), H. M. Powell, Edwin Garnet Stanford, Alice Christine Stickland, Percy Woodland.

A lecture on " Electric discharge lamps " (with demonstration) was given by Mr. E. B. Sawyer.

---

17 *July* 1942

The third meeting of THE OPTICAL GROUP. At the Imperial College, London, S.W. 7, Dr. A. O. Rankine being in the Chair.

A demonstration entitled " A trial of the Gaviola test for astronomical mirrors " was given by T. J. Tulley, B.Sc.

A paper on " The uses of optical instruments in the Army, and their influence on the test methods and limits ", by Major E. Golding, R.A., was read and discussed.

A lecture, with experimental demonstrations, on " The interferometer in lens and prism manufacture " was given by F. Twyman, F.R.S.

REPORT OF COUNCIL FOR THE TEN MONTHS ENDED 31 DECEMBER 1941

INTRODUCTORY AND GENERAL

THE Council has recently decided that the year covered by its annual report shall end on the same date as that covered by the Honorary Treasurer's Report and the Annual Accounts and Balance Sheet. The present report, therefore, covers the ten months ended 31 December 1941, instead of the year ended 28 February 1942.

The period under review is to be regarded as one of continued progress and success, evidence of which is to be seen in the growing roll of membership, the satisfactory attendances at most of the science meetings, the cordial approval and support given to the Colour Group and the more recently formed Optical Group, the outside sales of the *Proceedings*, *Reports* and other publications, and the encouraging financial position of the Society. To bring about this result in the present unfavourable circumstances has demanded a strenuous and continuous effort on the part of everyone concerned, senior and junior alike. That progress is being made in several directions is largely due to the response of many of our members to the *Special Notice* which, as in the two previous years, was issued by the officers at the opening of the session in October. This notice emphasized the need for whole-hearted co-operation, particularly in bringing the advantages of membership to the notice of possible new Fellows and Student Members, and in bringing the *Proceedings* to the notice of possible new advertisers.

Since the end of September, 1941,* the Society has been the sole tenant of the upper part of 1 Lowther Gardens, Exhibition Road, London, S.W. 7. Extremely favourable terms of tenancy were offered by Dr. Evelyn Shaw on behalf of the Royal Commissioners for the Exhibition of 1851 and accepted by the Officers and Council on behalf of the Society; and the Society has far more adequate accommodation for its office, its library and its stock than it would be able otherwise to obtain for a comparable, or even much greater, financial outlay. We wish to place on record the cordial thanks of the Society to Dr. Shaw and the Royal Commissioners for their generosity. Advantage has been, and is still being, taken of the opportunity afforded by the change for a rearrangement of the offices and publications. Fire-guard duties have fallen very heavily on the small staff, especially since the only male junior clerk left in September, 1941. The major part has necessarily fallen to the Honorary Secretary for Business: to him, and also to Mrs. Jevons and Dr. and Mrs. R. W. B. Stephens, who have taken a share of these duties, the thanks of the Society are accorded.

The representatives of the Society on other Bodies are the same as those given in the Council's Reports for 1939–40 and 1940–41.

MEETINGS

An Annual General Meeting was held at the Science Museum on 25 July 1941 for the presentation and adoption of the Reports of the Council and the Honorary Treasurer and for the election of the Officers and Council for 1941–42. The new President, Dr. C. G. (now Sir Charles) Darwin, nominated his predecessor, Professor Allan Ferguson, as Acting President during his impending absence; Dr. Darwin returned from the United States of America a few weeks after the end of the period under review.

In regard to Science Meetings, as was mentioned in last year's report, there is a greater demand for the exposition and discussion of a single topic than for the reading of several disconnected papers at a single meeting. The series of "Lecture surveys" and informal discussions begun last year was continued at four meetings, all at the Imperial College, the subjects and lecturers being: *Gravity Meters* (Dr. J. McG. Bruckshaw), *The Magnetic Hysteresis Cycle and its Interpretation* (Professor L. F. Bates), *The Mechanical Strength of Glass* (Professor W. E. S. Turner), and *The Use of Photo-electric Rectifier-type Cells in Photometry* (Professor J. MacGregor-Morris). Most of these "lecture-surveys" were accompanied by demonstrations, and five of the series, including the two reported last year, have been so far published in the *Proceedings*. The average attendance at these meetings was 40.

The remaining three science meetings were devoted to the 25th Guthrie Lecture, the 12th Thomas Young Oration and the 1st Charles Chree Address.

*** The Institute of Physics, whose office had been moved to the University of Reading on the outbreak of the war, continued to use 1 Lowther Gardens as its registered address until this date, when its tenancy of the rooms was terminated**

GUTHRIE LECTURE

The twenty-fifth Guthrie Lecture was delivered by Professor E. N. da C. Andrade on 4 April 1941 at the Royal Institution, the subject of the lecture being " A Problem of Guthrie's Time, The Sensitive Flame " (*Proceedings*, **53**, 329, 1941).

THOMAS YOUNG ORATION

The twelfth Thomas Young Oration was delivered on 30 May 1940 at the Science Museum by the Astronomer Royal, Dr. H. Spencer Jones, who took as his subject " The 200-inch Telescope " (*Proceedings*, **53**, 497, 1941).

CHARLES CHREE MEDAL, PRIZE AND ADDRESS

The award of the first Charles Chree Medal and Prize to Professor Sydney Chapman was announced in last year's Report. The presentation was made at the Science Museum on 25 July 1941, when Professor Chapman delivered the first Charles Chree Address on " (i) Charles Chree and his Work on Geomagnetism, (ii) Geomagnetic Time Relationships, and (iii) The Future of World Magnetic Surveying " (*Proceedings*, **53**, 629, 1941).

A suitably inscribed bronze replica of the medal was subsequently presented to Miss Jessie S. Chree, who was unable to be present when the Address was delivered.

DUDDELL MEDAL

The Council awarded the eighteenth Duddell Medal to Dr. William David Coolidge, of the Research Laboratories of the General Electric Company of America, Schenectady, New York, in recognition of his work in the development of new instruments and appliances, and especially of the heated-cathode high-vacuum x-ray tube. Since the end of the period covered by this report the medal has been presented to Dr. Coolidge by Viscount Halifax, H.M. Ambassador at Washington, at a meeting of the American Physical Society at Baltimore on 1 May 1942 (*Proceedings*, **54**, 209 and 527, 1942).

PUBLICATIONS

The two-monthly parts of the *Proceedings* have been issued as early as possible after the usual dates of publication. As was to be expected, Volume 53 (1941) is considerably smaller than Volume 52, the sizes of these being about 72 and 82 per cent, respectively, of the average size of the two preceding volumes. A further diminution will be necessary in Volume 54 (1942), as the supply of paper for the Society's publications is now (since the middle of 1941) issued under licence by the Ministry of Supply Paper Control, and the utmost care has to be taken to enable the work of the Society to be continued with the much reduced quantities of paper allotted at present, and the still smaller quantities licensed for the near future. The September issue of the *Proceedings* was the first to be printed on the new paper, which, though inferior to our former stock (purchased in 1940), is still of good quality.

The space devoted to advertisements in the *Proceedings* has steadily increased, and has now almost reached the maximum annual total permitted by the Paper Control. Throughout 1941 space was again put at the disposal of the Red Cross and St. John War Organization and the National Savings Committee.

Volume 7 (1940) of the *Reports on Progress in Physics* was published in April 1941, and Volume 8 (1941), the preparation of which was begun early in the period under review, has now been published (April 1942). The sales of Volume 7 have been very satisfactory, and the demand for earlier volumes of the series has been well maintained. In order to meet this demand the Society still wishes to purchase copies of Volumes 1 (1934), 2 (1935), and 4 (1937).

During 1941, the Society, in common with all publishing bodies in this country, experienced increasing difficulties caused by the additions to the long list of countries to which scientific publications may not be exported. We are fortunate in having an increased number of subscribers to the *Proceedings* both in the U.S.A. and in the U.S.S.R. ; but nearly all other overseas markets outside the British Empire are now closed to us. Losses of copies of our publications in ocean transit have, of course, been considerable at certain times. The Council again records its thanks to Miss Madeline M. Mitchell, Publications Manager of the American Institute of Physics, for her continued kindness in making the volumes of *Reports on Progress in Physics* available at the office of the Institute, 175 Fifth Avenue, New York, and in bringing the publications and membership of our Society to the notice of American physicists.

MEMBERSHIP

From the following table of membership figures at the end of 1941 it is gratifying to note (i) that the slight decrease in the total reported last year has been far more than made up; (ii) that each of the two classes of membership has increased; and (iii) that of those who had reached the end of their period of Student Membership more than two-thirds were transferred to Fellowship. It should be added that the December total of 1107, which was the highest in the history of the Society, has since been improved upon.

MEMBERSHIP ROLL ON 31 DECEMBER 1941

		Honorary Fellows	Honorary Fellows, Optical Society	Ex-officio Fellows	Fellows	Student Members	Total
31 Dec. 1940, Total		11	3	4	935	117	1070
Changes during 1941	Elected				28 } +41	51	
	Transferred				13	13 } −20	
	Deceased				9 } −35	2	
	Resigned or lapsed				20	3	
	Suspended				6	2	
	Net increase				+6	+31	+37
31 Dec. 1941, Total		11	3	4	941	148	1107

OBITUARY

The Council records with regret the deaths of the following Fellows :—Dr. R. T. Beatty, Dr. M. Benjamin, Mr. I. O. Griffith, Mr. W. H. Heaton, Dr. G. W. C. Kaye, Dr. C. G. Schoneboom, Mr. J. J. Stewart, Mr. R. J. Trump, and also the deaths of Mr. R. A. Day and Mr. J. Hutchinson, Student Members of the Society.

COLOUR GROUP

Five scientific meetings of the Colour Group were held in 1941, four in London and one in Bradford, the average attendance being about 50. The subjects discussed and other particulars are as follows :—

(i) *Colour Tolerance* ; at the Polytechnic, Regent Street, on 12 February ; an introductory paper by Dr. W. D. Wright and a group of three papers, two of which, by Mr. J. W. Perry and Mr. H. W. Ellis, on the physical and technical aspects, respectively, have been published in the *Proceedings*, **53**, 273 and 288 (1941). A short account of this inaugural meeting, prepared by Dr. Wright, was published in *Nature*, **147**, 303 (8 March 1941).

(ii) *Colour Terminology* ; at the Royal Photographic Society, on 30 April ; a paper by Mr. H. D. Murray, published in the *Journal of the Oil and Colour Chemists' Association*, **25**, 205 (1941).

(iii) *Colour Mixture Problems in the Dyeing Industry* ; at Bradford, on 25 June ; a lecture by Mr. J. G. Grundy, published in the *Proceedings*, **54**, 1 (1942).

(iv) *Whiteness* ; at the Imperial College, on 24 September ; a group of three papers by Mr. J. G. Holmes, Dr. V. G. W. Harrison and Mr. C. G. Heys Hallett, respectively, published in the *Proceedings* **54**, 81, 86 and 98 (1942).

(v) *Colour Blindness and its Importance in Relation to Industry* ; at the Electric Lamp Manufacturers' Association Lighting Service Bureau, on 10 December ; paper by Dr. F. H. G. Pitt, published in the *Proceedings*, **54**, 219 (1942).

The supply of off-prints of the papers and discussions to members of the Group has been arranged for. Arrangements have also been made for the exchange of off-prints with the Inter-Society Colour Council in the U.S.A., with whom close contact has been maintained.

In accordance with the item of the constitution which provides for the participation of certain other Societies, Institutions and Associations in the organization and activities of the Group, the following are now enlisted as participating Bodies : British Kinematograph Society, Illuminating Engineering Society, Institute of Physics, Royal Photographic Society, Society of Dyers and Colourists, Society of Glass Technology, Oil and Colour Chemists' Association, Physiological Society, and certain Committees of the British Standards Institution. It is evident from this list that the Group has already succeeded in bringing many colour interests together, a result of no small value to both science and industry.

Two sub-committees have been set up, to deal with Colour Terminology and with Colour Blindness in Industry, respectively.

The membership of the Group on 31 December 1941 was made up as follows :—

Members of the Physical Society	47
Members of other participating bodies	63
Members of six firms subscribing for sustaining membership ..	17
Other individual members	5
Total ..	132

OPTICAL GROUP

For some time past the officers of the Society have received many suggestions that a Group should be formed under the auspices of the Society, on lines similar to those of the Colour Group, for the consideration of scientific and technical problems in Optics. It was felt that, by holding meetings for the discussion of matters of current interest, the proposed Group could perform an important and valuable service in the present emergency, in which so many of those engaged on optical work are experiencing unexpected difficulties in the solution of their new problems.

The proposed Group should have the much desired effect of bringing together again many who have felt that their particular needs have not been met adequately since the merging of the Optical Society and the Physical Society of London. Indeed, it would appear that such a Group is the best, or even the only, means whereby the enlarged Physical Society can meet those needs.

The formation of the Optical Group was recommended at a well-attended preliminary meeting of those interested on 17 December 1941, and approved by the Council on 23 December, and the following were nominated as its first officers and committee :

Chairman : Dr. A. O. Rankine.

Honorary Secretary : Professor L. C. Martin.

Committee : Instr.-Capt. T. Y. Baker, Mr. R. J. Bracey, Mr. W. H. A. Fincham, Dr. V. G. W. Harrison, Mr. W. C. Hynd, Dr. H. Lowery, Capt. T. Martin, Mr. J. Perry, Mr. E. W. H. Selwyn.

Fellows and Student Members of the Society are eligible for membership of the Optical Group without additional annual subscription ; members of several other Societies, Institutions and Associations are eligible upon payment of an annual subscription of 5*s.* ; and other persons are eligible upon introduction by a member of the Physical Society or of one of these participating bodies and payment of an annual subscription of 10*s.* 6*d.*

It is proposed that about six meetings of the Group shall be held in each year. Papers read and discussed at these meetings may be communicated for publication in the *Proceedings* or in any other appropriate journal. A summary of the papers and a record of the activities of the Group will be prepared annually by the Chairman and published in the *Proceedings*.

REPORT OF THE HONORARY TREASURER FOR THE YEAR ENDED 31 DECEMBER 1941

The accounts show an excess of income over expenditure in 1941 of about £140. This result was achieved, however, only with the aid of a publications grant of £200 by the Royal Society from the Rockefeller Foundation Gift, which is gratefully acknowledged. The financial position at the end of the year was thus not quite so favourable as at the end of 1940, when the excess of income (apart from a very generous anonymous gift) was about £192. In the expenditure account every item but one shows a small increase: the increase in connection with the tenancy of 1 Lowther Gardens is accounted for partly by Air Raid Precautions and partly by the termination of the tenancy of the Institute of Physics at the end of the third quarter—the Physical Society now being responsible for the whole of this expense. In the income account it is satisfactory to note appreciable increases in the entrance fees, annual subscriptions, sales of publications, and advertisement revenue.

No change has been made in the Society's investments, the total value of which on 31 December 1941, as given in a statement by the Manager of the Charing Cross Branch of Westminster Bank, Ltd., was £620 higher than at the end of 1940. The stock of publications, which has not been re-valued, has changed but little in total amount during the year.

(Signed) C. C. PATERSON,
Honorary Treasurer.

18 *August* 1942.

INCOME AND EXPENDITURE ACCOUNT FOR THE YEAR ENDED 31 DECEMBER 1941

1940 £	EXPENDITURE	£	s.	d.	£	s.	d.
	To						
95	" Science Abstracts "				201	7	3
	Normal Publications :						
1492	" Proceedings "	1905	4	11			
52	Agenda papers and Notices	37	19	3			
674	" Reports on Progress in Physics "	693	0	0			
					2636	4	2
286	Postage on Normal Publications and General Correspondence				234	6	6
26	Expenses at Meetings				43	15	10
—	Honoraria to Special Lecturers				40	0	0
	Administration Expenses :						
754	Secretarial and Clerical Assistance and Office Expenses	760	7	5			
74	1 Lowther Gardens (including Air Raid Precautions)	95	6	0			
					855	13	5
892	*Balance, being excess of Income over Expenditure, carried forward to Accumulated Fund*				140	1	1
£4345					£4151	8	3

1940 £	INCOME	£	s.	d.	£	s.	d.
	By						
	Subscriptions :						
19	Entrance Fees	35	14	0			
1655	Fellows	1691	13	8			
58	Student Members	85	1	6			
—	Colour Group	23	8	6			
105	For " Science Abstracts " and Advance Proofs	111	19	4			
					1947	17	0
	Sales :						
649	" Proceedings "	797	7	8			
707	" Reports on Progress in Physics "	616	10	3			
4	Special Publications	47	16	9			
					1461	14	8
100	Advertisements in " Proceedings "				235	19	10
278	Dividends from Investments				285	16	9
70	Transfer from Life Composition Fund of amounts paid by Fellows now deceased				20	0	0
—	Publications Grant by the Royal Society from the Rockefeller Foundation Gift				200	0	0
700	Anonymous Donations				–	–	–
£4345					£4151	8	3

BALANCE SHEET AS ON 31 DECEMBER 1941

LIABILITIES	£	s.	d.	£	s.	d.
Sundry Creditors				120	3	9
Life Compositions :						
As on 31 December 1940	1501	0	0			
Less Transfer to Income and Expenditure Account	20	0	0			
				1481	0	0
Subscriptions received in advance :						
Members	43	19	2			
Publications	51	2	1			
				95	1	3
SPECIAL FUNDS :						
W. F. Stanley Trust Fund	259	0	0			
Duddell Memorial Trust Fund	379	4	9			
"Progress Reports" Reserve Account	83	1	0			
Herbert Spencer Legacy	239	7	6			
Charles Chree Medal and Prize Fund	1880	12	0			
				2841	5	3
GENERAL FUND :						
As on 31 December 1940	4206	18	9			
Add Balance of Income and Expenditure Account	140	1	1			
				4346	19	10
				£8884	10	1

ASSETS	£	s.	d.	£	s.	d.
Investments at Market Value on 31 December 1939 :						
W. F. Stanley Trust Fund	259	0	0			
Duddell Memorial Fund	374	0	0			
Charles Chree Medal and Prize Fund	1866	10	0			
Physical Society	5917	10	0			
				8417	0	0
(Market Value on 31 December 1941 : £9650.)						
Dividends due from Investments				54	12	1
Inland Revenue—Income Tax recoverable for 1941				55	10	10
Subscriptions due				75	15	10
Sundry Debtors				135	6	5
Stock of Paper and Binding Material				95	8	4
Cash at Bank	35	12	4			
,, *in Hand*	14	10	11			
,, *in Post Office Savings Bank Account*		13	4			
				50	16	7
				£8884	10	1

We have audited the above Balance Sheet and have obtained all the information and explanations we have required. We have verified the bank balances and the Investments. In our opinion such Balance Sheet is properly drawn up so as to exhibit a true and correct view of the state of the Society's affairs according to the best of our information and the explanations given to us and as shown by the books of the Society.

SPENCER HOUSE, SOUTH PLACE, E.C. 2.
28th August 1942.

KNOX, CROPPER & Co.,
Chartered Accountants.

LIFE COMPOSITION FUND ON 31 DECEMBER 1941

	£	s.	d.
29 Fellows paid £10	290	0	0
1 Fellow paid £15	15	0	0
14 Fellows paid £21	294	0	0
28 Fellows paid £31 10*s.*	882	0	0
	£1481	0	0

SPECIAL FUNDS

W. F. STANLEY TRUST FUND

	£	s.	d.		£	s.	d.
Carried to Balance Sheet	259	0	0	£300 Southern Railway Preferred Ordinary Stock	199	0	0
				£442 Southern Railway Deferred Ordinary Stock	60	0	0
	£259	0	0		£259	0	0

DUDDELL MEMORIAL TRUST FUND

Capital

	£	s.	d.		£	s.	d.
Carried to Balance Sheet	374	0	0	£400 3½% War Loan Inscribed "B" Account	374	0	0

Revenue

	£	s.	d.		£	s.	d.
Balance on 31 December 1940	8	15	3	Interest on War Loan	14	0	0
Balance carried to Balance Sheet	5	4	9				
	£14	0	0		£14	0	0

"PROGRESS REPORTS" RESERVE ACCOUNT

	£	s.	d.		£	s.	d.
Balance carried to Balance Sheet	83	1	0	Balance on 31 December 1940	83	1	0

HERBERT SPENCER LEGACY

	£	s.	d.		£	s.	d.
Balance carried to Balance Sheet	239	7	6	Balance on 31 December 1940	239	7	6

CHARLES CHREE MEDAL AND PRIZE FUND

Capital

	£	s.	d.		£	s.	d.
Design of Medal and Preparation of Die	200	0	0	Balance on 31 December 1940	2065	16	4
Balance carried to Balance Sheet	1865	16	4				
	£2065	16	4		£2065	16	4

Revenue

	£	s.	d.		£	s.	d.
Charles Chree Prize	136	10	0	Balance on 31 December 1940	106	9	0
,, ,, Medal	4	14	6	Interest on Investments	68	17	2
Printing Certificates	19	6	0				
Balance carried to Balance Sheet	14	15	8				
	£175	6	2		£175	6	2

[*Photo*: *Lafayette*.

PROF. ALLAN FERGUSON,
President of the Physical Society, 1938–41.

[*Reproduced by permission from Quart. J. R. Met. Soc., Vol.* 60 (1934).

PROF. SYDNEY CHAPMAN, M.A., D.Sc., F.R.S.
First (1941) Chree Medallist

THE PROCEEDINGS OF
THE PHYSICAL SOCIETY

VOL. 54, PART 1 1 January 1942 No. 301

14. APR. 1942
LIBRARY

PROBLEMS OF COLOUR MIXING IN THE DYEING INDUSTRY

BY J. G. GRUNDY,

Lecture delivered to the Colour Group, 25 *June* 1941

FOREWORD

THIS lecture has been compiled to indicate to the physicist the trend of thought of the practical dyer in his choice of dyestuffs. The lecture contains nothing new for the dyeing industry.

It must be understood that many of the terms used in the dyeing and textile industries have not necessarily the same meaning as when used by the physicist. To give an example, whenever the term " shade " is used in this paper, it is equivalent to the term " hue " as used by the physicist. Although there is a considerable looseness in the dyeing industry as regards terminology and expressions, it is not within the scope of this paper to clarify this position. This looseness in expressions can be indicated as follows:—

If two dyeings appear *similar* in all respects they may be described as accurate match, dead match, on shade or up to pattern. If a pattern apparently contains *less* dyestuff than another it is described as weaker, thinner, greedy, paler, empty, lacking in depth, lighter, starved, skimpy or " wanny ", whereas if a pattern apparently contains *more* dye than another it is described as too intense, heavier, fuller, stronger, darker, deeper, denser, overloaded, bursting or " brussen " with colour. All these terms are in common use in the dyeing trade.

[The lecturer gave a historical account of the art of dyeing from the earliest Chinese and Indian work, and the almost contemporaneous Egyptian art, specimens of which are found in the Valley of Tombs, to modern times. It is regretted that space does not permit the inclusion of this account, which mentioned, *inter alia*, that in the time of Pliny indigo was known only as a pigment and not as a dye, but that he knew of the use by the Egyptians of the polygenetic type of dyeing.—ED.]

The present-day dyer is rarely concerned with natural colouring matters and pigments. The introduction of synthetic dyestuffs in 1885, and the rapid progress by the dyestuff manufacturers in the production of dyes, revolutionized dyeing procedure.

In the years just previous to the present war the world production of synthetic dyestuffs amounted to seven hundred million pounds weight annually, and most of this poundage was consumed by the textile industries. The Colour Index, which is the standard English dictionary on dyestuffs, describes over fourteen hundred distinct chemical compounds, nearly all of which find application in the textile and allied trades. It can be taken that most of these products are in everyday use. It will be shown later that the endeavours of the organic chemist to extend the ranges of dyestuffs are fully justified.

As reference will be made later to " primary colours ", it should be explained that, from a textile standpoint, as yellow cannot be produced by mixing any other dyes together, it has to be looked upon as a primary colour, and, therefore, in the dyeing and allied industries, red, yellow and blue are accepted as the three standard primaries. Owing to the wide differences in shade, fastness properties and methods of application of the dyestuffs available, however, the modern dyer has really to consider not three primaries, red, yellow and blue, but actually dozens of reds, yellows and blues.

Those not intimate with the industry will wonder why there should be the necessity for such a wide variety of red, orange, yellow, blue, brown, green and black dyes, when in theory one can produce every possible shade with the three primary colours, red, yellow and blue. There are several reasons for this. The selection of a dye is not dependent upon shade only, but on the type of material to be dyed and the fastness properties required. Dyes suitable for the dyeing of one fibre are not necessarily suitable for the dyeing of other fibres.

This point can be illustrated by Kiton Fast Green V on a cotton and wool union cloth. This colour dyes the wool and leaves the cotton undyed. A further example can be instanced: Chlorantine Fast Red K when applied on viscose rayon and acetate rayon, dyes the viscose rayon red and leaves the acetate rayon undyed. Again, dyes used for cotton are, with few exceptions, not applicable in commercial practice for the dyeing of wool, and only in very rare cases are dyes for wool used in cotton dyeing. Moreover, the dyeing procedure in the two industries is quite different. Only in a few branches of the industry are shades built up on the basis of red, yellow and blue. A large number of the shades dyed for the ladies' dress goods trade, and hosiery and carpet yarns, are obtained by using combinations of red, yellow and blue. The restriction in the use of red, yellow and blue is due partly to the availability of a large number of homogeneous dyes which yield orange, brown, green, violet and black shades. In general, these homogeneous dyes form a much more suitable basis for shades which are of the same character than combinations of the three extremes, red, yellow and blue. The shades obtained with the homogeneous oranges, browns, greens, violets and blacks can be adjusted with dyes of a similar tinctorial character, or with the primary colours, red, yellow and blue, as may be required. It should also be understood that as the reds, yellows and blues are themselves not pure in character, there are orange, violet and green dyes manufactured which cannot be matched for purity in shade with combinations of the purest primaries, red, yellow and blue.

This point can be clearly demonstrated by comparing a dyeing of Kiton Fast Green V on wool with the nearest shade of green to this obtained with a combination of a bright yellow and a bright blue dye. The combination of the two latter dyes yields shades very dull compared with the dyeing of Kiton Fast Green V.

It is essential for the dyestuff to possess fastness properties to meet the requirements of the dyed material from the point of view of subsequent finishing and making-up processes, and the expectations regarding the general utility of the finished garments.

[The lecturer then gave a detailed account of the fastness properties of various classes of dyes and continued :]

The dye manufacturers classify their dyestuffs into groups. For example, for cotton and regenerated celluloses of the viscose type, the groups are:—

1. Basic dyes.
2. Direct dyes, which are sub-divided as follows:
    (*a*) Direct dyes of average fastness properties.
    (*b*) Direct dyes fast to light.
    (*c*) Direct dyes which can be after-treated:
        (i) With formaldehyde to increase fastness to washing.
        (ii) With chrome and copper salts to increase fastness to washing and light.
        (iii) Those dyestuffs which can be diazotized on the fibre and developed with suitable naphthols or diamines, thereby producing shades fast to washing.
3. Sulphur dyes.
4. Azoic dyes.
5. Vat dyes.

For wool dyeing, the groups are:—

1. Acid dyeing dyes.
2. Neutral dyeing dyes.
3. The Neolan class of dyes.
4. Mordant dyes.

It will be appreciated that very few dyers have more than one trade for which to cater. For example, dyers in the cotton trade are dyers of loose cotton, cotton in hank, cheese or cop form, cotton piece goods or cotton knitted goods. In the case of the wool trade, dyers are either loose wool and slubbing dyers, or wool yarn or piece dyers. In addition, of course, there are many special branches of the dyeing industry, such as silk dyers and dyers of hose and rayons. Mention must be made of the fact that most dye houses have to consider the dyeing of unions in one form or another along with their normal run of trade.

Having probably only one class of dyestuff to consider, or, if more than one, closely allied classes, on account of the dyeing properties and the economics of dyeing, the dyer must familiarize himself with the dyeing and fastness properties of the range of colours with which he is immediately concerned. He is, therefore, readily able to make a selection of dyestuffs for a specific job.

A consideration of the available dyes on the market for any particular branch of the dyeing industry is rather interesting. For example, in cotton dyeing there are at least sixty-two distinctly different red dyestuffs in use, as the firm with which we are associated actually market this number.

They can be classified as follows, according to the more important properties which are required in the various trades connected with cotton dyeing:—

Total number of dyes	62
Water-soluble dyes	43
Dyes insoluble in water	19
Dyes fast to light, five or over	20
Dyes fast to chlorine	17
Dyes fast to washing	26
Dyes dischargeable with hydrosulphite	30
Dyes which reserve acetate cellulose rayon	13
Dyes fast to light and washing	17
Dyes fast to light, washing and chlorine	16
Dyes fast to light and dischargeable	10
Dyes fast to washing and dischargeable	15
Dyes fast to light, washing and dischargeable	7

[The lecturer then considered the differing requirements as regards fastness, according to the use to which cloth is to be put. Thus, men's hats, with their longer life and probable dry-cleaning, present different problems from women's hats. Suitings, linings, stockings, bathing costumes, curtains, carpets in houses, in hotels and on ships, and upholstery, all make different demands on the fastness to washing, to light, to perspiration, to sea-water, and so on. The lecturer then proceeded as below.—Ed.]

It will be appreciated that when selecting a dye for a specific purpose where certain fastness properties are required in the finished goods, the number of dyes suitable is considerably reduced. It is essential, of course, that the dye selected should give the correct shade, and this factor itself eliminates a large proportion of the available products.

There is no known method of determining by photometric or other scientific means how much colour will be required to obtain a match to a pattern and, if a combination of colours is required, what quantities of each colour. These factors have to be pre-determined by laboratory dyeing experiments or gained by experience in large-scale practice. In view of the fact that the laboratory dyeing is carried out in the same way as in large-scale practice, dyeing can be considered purely empirical.

These remarks are supported by the following:—

A. The solution of a dyestuff in process of application in many instances bears no resemblance to the shade of the finished dyeings. This is especially the case with dyes known as sulphur, azoic, vat and chrome colours.

B. The final shade with many dyes is only developed by chemical processes operated in the actual dye-bath. The most important class of dyes in this category are the chrome fast colours, so called since they require chromium salts or chromates for the development of their shades. Without this treatment the shades produced by dyes of this group are, in general, of no commercial value, some being almost colourless.

C. The dyeing properties of individual dyestuffs irrespective of their shades or fastness properties vary considerably in their rate of dyeing and the degree of exhaustion under similar dyeing conditions. Therefore, in compound shades, when two or more dyestuffs are used, the number of variables is considerably increased.

D. Again, the physical appearance of the fibres to be dyed, and the structure of the material, play an important part in the resultant shades. For example, 2 per cent shades of the same dyestuff on two qualities of viscose rayon yield entirely different results. An extreme difference is obtained if the dyeings are carried out on filament and spun viscose, or on matt and bright acetate rayon.

E. Union fabrics containing two or more classes of fibres may be required to be dyed a full shade or a multi-coloured effect. It is quite common to have six to nine distinct dyestuffs in the one dye-bath.

The usual procedure of the dyer is to set out his day's work and to give instructions to the workmen in the dyehouse for the charging of the dye-baths. One dyer will control from 20 to 40 dyeing machines at the one time. His instructions are carried out by dyehouse labourers under the supervision of a foreman, depending upon the class of goods processed. As the dyeing proceeds, and reaches a certain stage, the workmen take samples from the bulk dyeings to the dyer for control. The dyer, after comparing the sample with the pattern he has to match, gives the workmen such instructions as may be necessary to adjust the shade. This procedure is continued until a satisfactory result is obtained. The dyer's match is nearly always subjected to a further examination by an independent member of the works staff, usually someone employed in the making-up room, before despatching to the customer. Even under this dual control it is not infrequent to have goods returned from the warehouse of the merchant or from the makers-up on account of inaccuracy in matching.

It is obviously essential, therefore, that people associated with the dyeing industry should be normal as far as colour vision is concerned. Normality of colour vision has for some time been tested and checked by the use of pseudo-isochromatic plates, first produced by Stilling. In more recent years Ishihara in Japan has also designed a series of plates to test for colour blindness. Within recent months the American Optical Company has prepared a further set of plates, enlarging upon the ideas of Stilling and Ishihara. These plates have proved to be very reliable as far as concerns the dyeing trade.

Different textiles are not all viewed in the same way when matching. Some patterns are viewed only by transmitted light, others by reflected light, whilst in some branches of industry both these factors have to be taken into con-

I. A. R. I.

sideration. In the carpet and allied industries, where the cross-section of the fibre is the most obvious factor, dyeings under judgment are made into tufts, cut across to make a pompon, and compared on the brush-like surface. The intensity of the shade on the cut surface is much greater than on the yarn.

Since matching shades is relative rather than absolute, variations under different lighting conditions are experienced. Moreover, many patterns have to be matched on entirely different materials, using distinctly different chemical groups of dyestuffs. Having in mind such points, it is not surprising that disagreements arise as to what is a match, even amongst people who can be considered to have normal colour vision. This is particularly marked when dyeings are examined at different times, as in most cases daylight is taken as the standard source of illumination, and this is by no means a constant. Whenever possible, therefore, comparisons should be carried out on similar materials and under identical conditions.

Attempts have been made in recent years to arrive at some defined standard of artificial illumination with a view to eliminating the difficulties which arise in matching patterns under the varying conditions of daylight. In spite of a large amount of work which has been carried out in this direction, daylight is still generally accepted, a good north light being taken for preference.

It must not be overlooked that many shades of fabrics are used for evening wear, and chosen, not for their tone when examined in daylight, but for their tone under actual conditions of use, that is, in artificial light. For example, dance frocks, materials for stage wear, men's dinner and evening wear, must be considered mainly from this point of view. These factors, therefore, are specially watched by the dyer when dyeing and matching materials for these trades.

It has been stated that normal colour vision is only one-tenth as critical as known scientific instruments, and probably this statement is no exaggeration. A higher standard than normal vision for matching, however, would not be practical.

Matching to pattern in the textile trade is of high standard, but for reasons which have been described, a certain amount of tolerance must be allowed, otherwise few dye-houses could be run for profit.

From the foregoing notes the following conclusions can be drawn:—

1. Dyeing is purely empirical.
2. It does not appear possible to devise any type of scientific apparatus which could make due allowances for all the variables which occur in dyeing, and which could predict with even reasonable accuracy the percentage of dye or dyes to use for the production of a specific shade.
3. Instruments of the colorimetric type would be of little use, as large numbers of dyes in everyday use are quite insoluble in water. Others give almost colourless solutions.
4. Again, as demonstrated with a few chrome dyes on wool, the colour of

the solution bears little or no resemblance to the final shade obtained; and there are many similar classes of dyestuff which require some form of after-treatment to develop the correct shade.

5. In combination shades of two or more dyestuffs, it cannot be assumed that the dyes exhaust equally under the different dyeing conditions, and this factor alone would increase the difficulties of using some form of scientific instrument for measuring the amounts of colour required.

6. From an examination of the dyeings exhibited it will be seen that the shade obtained by any one dye is dependent to a considerable extent on the fibre itself. Similar percentages of one dye can yield weaker or stronger shades on different grades and qualities of cotton, wool, artificial silk, etc., according to the physical appearance of the fibre.

7. In a great number of cases it is not a single class of fibre which has to be dyed, but a mixture of fibres. This is particularly the case under present-day conditions, where there is a definite shortage of wool and cotton, and where artificial fibres are being used along with wool and cotton in ever increasing quantities. It is not uncommon in dyeings of this type to use six to eight different dyes on the union. This, again, rules out any possibility of devising a scientific method of predicting with accuracy the amounts of dye to use.

It will be readily seen from these conclusions that the modern dyer has to be a man of skill and resource if he is to cope with the large number of varying factors which can arise in connection with his work.

We hope this very brief outline of the dyeing industry has proved of interest, and that it will be agreed that theoretical consideration in regard to the quality and proportion of dyes for the production of dyed goods can only be of passing interest.

You may rest assured, however, that if by some means or other a way can be devised to make the work of the dyer easier, more efficient and effective, such a discovery will be more than welcome in an industry which contributes to such a large extent to the pleasures of the human race throughout the world.

In closing, I must express my indebtedness to The Clayton Aniline Company, Limited, for permission to read this paper, and the appreciable help rendered by Messrs. F. Parrott and B. Kramrisch in compiling it.

DISCUSSION

Dr. Cassie (Wool Industries Research Association). A great obstacle to appreciation of colorimetry by industry is the lack of any straightforward account of the subject. Much has been written and most of it is unintelligible even to physicists, unless they are specialists in the subject. The principles of colorimetry are now well established: it should be possible to state these clearly, and give logical deductions of the details necessary for practical work. When such an account is available, co-operation between physicists and colourists will become possible.

Mr. **Warburton** (Wool Industries Research Association). A remark in the discussion about people seeing colours differently as they age is important, as limiting the useful accuracy in matching samples dyed with different types of dye, since no two people have the same degree of retinal pigmentation. This is shown by the fact that although normal persons get very nearly the same co-ordinates for any colour when using a colorimeter—as a result of the convention that equal quantities of the primaries match white—it is very rare for the actual instrument settings to be the same for two people, even under identical conditions. Similarly, the proportions in which three primary dyes will be required to match a given sample (dyed with different dyes) will vary slightly from person to person, in much the same way as will the proportions required to obtain a match in different types of lighting, although the variation will be much less. A perfect match for all observers can thus only be obtained by dyes having similar absorption spectra.

The lack of sensitivity frequently complained of in connection with instruments such as the Donaldson colorimeter could, in many cases, be overcome by the use of a larger field without detriment to the accuracy of the instrument.

Mr. J. W. **Perry**. We owe our thanks to Mr. Grundy for the considerable care and thought evidenced by his paper, demonstrations and experiments, and it is clear that the problems of the dye-chemist and dyer are both varied and profound. From present indications I do not think it is too much to hope that dyeing may one day become an exact science, but it is clear that at present, in spite of much theoretical work on the subject, dyeing is an art. With the growth of knowledge, however, what is an art may, given certain conditions, become a science. In the dyehouse the control of the dyeing process is based largely upon the application of a form of trained instinct, and in this are applied, consciously or subconsciously, facts and experience which, if properly analysed and co-ordinated, would themselves form the substance of a not inconsiderable body of knowledge. Such a science, developed by research, would bring a comprehension of the real nature of the problems and would also facilitate the control of the process by making clear its essential nature.

The various interesting facts and phenomena described and illustrated by Mr. Grundy certainly indicate that there is considerable scope for such research which would so assist the dyer. It would be unnecessary and, indeed, impossible to draw a precise and fine distinction between chemistry and physics in their application to such a subject; in fact, the whole chemical aspect of the problem may be regarded also as a physical one. The older ideas of the dyeing process as depending upon solid solution or colloidal action may yet prove to be fruitful, but in view of the very wide variation in the nature of the phenomena involved, more ample scope for their explanation is to be sought in the recent views favouring chemical, physical or electrostatic adsorption. The physical liaison through the adsorption bond passes imperceptibly, with changing circumstances, into chemical combination, as is apparently also the case in dyeing.

As is indicated by recent work [Ruggli and Jensen, *Helv. chim. Acta*, **18**, 624 (1935); Zechmeister, Cholnoky *et al.*, *Chromatography*, London, 1941; and Brode and Brooks, *J. Amer. Chem. Soc.* **63**, 923 (1941)], valuable information may be obtained from the application of adsorptive analysis to dye compounds, using the methods of the Tswett chromatogram [Tswett, *Bot. Z.* **63**, 273 (1905)], and it may be that the behaviour of the dye in relation to the textile fibre in simple cases is itself subject to similar conditions and interpretations. Consideration of the phenomenon of dyeing from the point of view of the ζ-potential and by the aid of electrophoresis should also throw some light upon its essential nature, and assist the fundamental development of the subject. Some dyes would also no doubt require further special investigation owing to peculiarities in their behaviour, such as dichroism, polarization, etc., and consideration of these in classes related to their molecular structure is likely to yield information of great importance. Quite apart, however, from the more far-reaching value of such researches, it is probably not too much to say that specific differential spectrophotometric characteristics, by which the colour could be controlled and which could readily be found experimentally and tabulated in relation to certain conditions, would apply in all cases. In any given instance, therefore, given the necessary data (which could be obtained from previously recorded tests), the ultimate result could be reached with precision by the application of a definite spectro-photometric procedure, provided the essential conditions were known and controlled.

These remarks refer mainly to the chemical and physico-chemical aspects of the unsolved problems presented by Mr. Grundy, but there have also been important problems of a physical nature, and although it is not suggested that these are all completely solved, especially, as pointed out by Mr. White, in regard to the analysis of the complete reaction of the dyed fabric to incident light, yet it is important that those in the industries interested in colour should realize the great forward strides which have been made in this direction during the last quarter of a century. It is, indeed, unlikely that comparable progress has been made during the same period in any associated field. The physicist has here gone a good third of the way toward solving the dyer's problem, in that, out of a confused mass of unco-ordinated data and observations, the basic relationship of colour, as a visual function, to light, a form of energy, has been firmly established and formulated in generally acceptable terms. Thus the uncertainties of what is sometimes called the " psycho-physics " of colour have been removed from the arena, international agreement has been reached upon a system and an acceptable scale for designating colour in numerical terms and accurate colorimeters have been introduced, based upon the new system. This must, sooner or later, inevitably call for revision of outlook by all those interested in colour and colouring materials, for it provides the method and means for accurately specifying the main and ultimate purpose of the industries involved.

Colorimetric analysis has certainly required many years for its emergence and standardization, however, and the dyer may perhaps be excused for failing to take it seriously until it affects him more nearly. Moreover, as indicated by Mr. Grundy, the dyer's colour-problem is really somewhat broader than is comprehended within the difficulties of colour-matching.

It was early recognized that colour is not directly related to physical constitution, but that it contains an important clue thereto; further, that if the chemist will think in terms, not of colour alone, but also of the absorption of light, important generalizations are made possible. This fact, which relates chemistry to spectrophotometry, is among those which, as Mr. Bunbury has remarked, have given the dye-chemist considerable power over the production of colour, representing a substantial and respectable achievement.

Equally the dyer may profit, now that the relationship of colour to physical stimulus has been firmly established. His problem, now less extensive and more clear-cut, may be regarded as consisting of (*a*) the production of substances having the requisite chemical, physical and tinctorial properties in relation to certain materials, and (*b*) fixing them, with or without modification of molecular structure, to textile fibres, again with due regard to physical and chemical conditions, so that the nett result is the production of an accurately specifiable physical reaction of the fibre to light. His problem may be to produce either a colour match in a specified illuminant or a spectrophotometric match, which ensures a colour match in all illuminants. This is an important distinction and really contains one more condition, which may be added as a refinement to the already seemingly formidable list of circumstances and conditions to be taken account of by the dyer and instanced by Mr. Grundy, viz., a specified degree of approach to spectrophotometric match; for upon this depends both the variation in colour which will take place when the illumination upon the dyed material is changed in character, for example, say from that of daylight to that of artificial light; and also the apparent differences in colour and matching properties, due to increased macular pigmentation in older observers. To achieve this condition in varying circumstances, however, precludes the possibility of using only three primary colours, whatever they may be, and it may be as well to emphasize that no definite number of colouring materials can be specified as being generally adequate to secure a reasonably exact spectrophotometric match, although in any given case it would require only a simple spectrophotometric test to ascertain what would enable the required conditions to be fulfilled.

May I add, in conclusion, that I think the wealth of problems still confronting the dyer, and very strikingly demonstrated by Mr. Grundy, afford ripe material for organized research, which I hope may some time be carried out, as well to benefit general knowledge, as also to relieve the perplexity of the dyer and to aid him in his work; but I hope also that the above remarks in appreciation of Mr. Grundy's paper will contribute to make it clear to those interested in applying physical methods to dyeing that the advances which have already been made pave the way for a reconsideration of the dyer's art on modern lines and enable the processes which he uses to be more carefully analysed and controlled.

In regard to the question raised as to a suitable method of measurement to enable the accurate reproduction of a " black "; this is a photometric rather than a colorimetric problem, as the insensitivity of the eye to chromaticity differences at very low brightness-levels causes chromaticity differences to be of subsidiary importance. The first step should, I think, be to agree upon the colour which will be called " black ", since at best this will have some brightness and may also not be neutral; it will, in fact, be, physically, a pseudo-black. The light received from such a colour for different regions of the spectrum under controlled conditions of illumination may be measured against a standard, derived from the same source or from another operating in series and at the same efficiency as that source, by means of a low-brightness photometer used in conjunction with calibrated selective colour filters.

Dr. W. D. Wright. Mr. Grundy has given a very excellent survey of the problems that confront the dyer, a survey that is especially helpful to the physicist with little knowledge of the industry. He has made it abundantly clear that while colour is the first consideration, other factors, especially those of fastness, may determine which of a number of dyes is to be preferred for any particular purpose. It is, however, surprising to the physicist that since dyeing aims primarily at the production of colour, Mr. Grundy has not found it necessary to refer to the properties of a dye that give it its colouring power, namely, the extent to which the various parts of the spectrum are absorbed when light strikes a dyed surface. This action, it is true, is a physical action, but it is none the less fundamental to the business of dyeing. It is, in fact, interesting to speculate on how the dyeing industry would have developed if no dyeing had ever been carried out until the last few years. Some business magnate might have approached a group of scientists with the idea of establishing an industry for applying colour to materials. The scientists would have reported that colouring involved the absorption of light and was essentially a physical operation. A physicist might therefore have been appointed director of research, and he would no doubt have collected a number of physicists and chemists to assist him. Asked to produce a given colour, he would have advised the measurement of the absorption curves of certain dyes, and by a suitable combination would have endeavoured to produce the desired result. Before very long it would have become apparent that the practical problems of dyeing were mainly chemical, and the research director might have been asked to accept a more junior position. Nevertheless, the conception of the colouring aspect of dyeing would have been continuously linked up with the absorption curves of the dyes, and continual efforts would have been made to develop improved apparatus for measuring these curves.

This is pure fantasy, but it helps in deciding whether the physicist can now be of any service in the dyeing industry under the conditions in which it has in practice developed, or whether he would merely be a nuisance. There seem to be two ways in which the physicist might be able to demonstrate his worth. In the first, if the dyer has any outstanding problems that he cannot solve, it may be that the physicist can help him to find the solution. In the second, it might be that if physicists undertook a long-term research into the colouring effects of dyes, they might be able to make fundamental recommendations that would revolutionize the art. This would certainly take time, and is purely speculative; but one would naturally expect the dyers to be extremely anxious to get to the fundamentals of their trade—and when they get there they will find physics. Unless they understand the physics of colour the result of their colour-mixing experiments must remain more or less of a mystery to them. It would be rash to claim that increased knowledge means increased dividends, but that is the frequent experience when scientific research is undertaken. I would, therefore, urge Mr. Grundy to consider whether research into the physics of colour as applied to dyeing ought not to be undertaken on a vastly greater scale than at present. I would not claim to have a detailed programme of research to offer, still less would I claim that all the apparatus needed for the research is available, but I do believe that an overwhelming case can be made out that the research should be attempted.

If a long-term programme is not approved, can the physicist prove his worth by helping with current problems? Mr. Grundy has referred to the case of a particular red dye that gives a more brilliant colour than any other. I would like to ask him whether he knows why this dye has this particular property? Does he believe that further chemical research is justified in trying to produce a still more brilliant colour of that hue, or has the maximum possible brilliance been obtained? Are there other colours whose brilliance might be increased, or is it the case that in general the most brilliant dyes possible are now available? These questions can be answered once the spectral absorption curves have been measured, but I do not know how else they can be settled. It is true that the curves may be of little direct use in helping to produce a new dye, but they will show whether time and money should be spent on the attempt, and they will show the properties the dye should, if possible, be given. I should be interested to know whether Mr. Grundy would agree with this point of view.

Another problem that appears to cause the dyer much trouble is that two samples may match in one light but not in another. This difficulty would disappear if the spectral reflection curves were of the same type—again, surely, a case for spectro-photometric measurements.

In the samples I showed at the meeting I raised the question of the colour that would result when the yellow and blue samples of slubbing were gilled together. As Mr. Grundy remarked, that is, in part at least, optical colour mixing, and from tests made before the meeting it certainly appeared that the physicist was able to prophesy the colour more accurately than the dyer. Yet this is a process in constant use in the dyeing industry—another indication that use could be made of the principles known to the physicist.

I hope this discussion will provide some evidence of the value of increased co-operation between the physicist, chemist and dyer. If the Colour Group helps to develop this co-operation it will, I believe, be doing a very useful service.

AUTHOR'S REPLY. The contributions to the discussion have, in many instances, given emphasis to the complexity of the dyer's task, whereas from the physicist some hope is given that the difficulties in regard to the dyer's choice of dyes and quantities required for a given shade may eventually be more easily and more accurately controlled by the use of instruments not yet devised No doubt this possibility opens up a vast field of research, and in this regard the physicist must not lose sight of the fact that methods of scientific control must compete, not only in accuracy but in speed, with the empirical methods at present used by the dyer.

The physicist must also bear in mind that dyestuffs are technical products, and the dyer may from time to time have to use the same dyestuff manufactured by different dyestuff-producing firms, which, when dyed, yield similar shades to one another, but in regard to physical properties, that is, appearance of powder and appearance in an aqueous solution, vary very considerably.

It is not correct to assume that dyeing is wholly an art developed by years of practice and experience. The process of modern dyeing is more accurately described as an art well supported by scientific control. I see no reason to support the idea suggested by Dr. Wright that the dyer's colour-mixing experience is a mystery to him. The dyer is very intimate with his knowledge of colour-mixing as far as it is applied to the dyeing industry.

In regard to the possibility of two samples being a match in one light and not in another, Dr. Wright may be under some misapprehension, having regard to the remarks which he made. It was intended to refer to patterns which had been dyed with combinations of dyestuffs rather than to the behaviour of two or more individual dyestuffs.

It should be emphasized that there is no desire on the part of the dyeing industry to raise the standard of matching. This is already high enough, and most people associated with the dyeing and textile industries will readily agree that a higher standard than is generally maintained would not be a practical and economic proposition.

The discussion would indicate that there is a very wide gap in the trend of thought between the dyer and the physicist, and no doubt the narrowing of this gap by closer collaboration and understanding of extremes in point of view will be of considerable value towards the attainment of that goal which is the aim of the physicist.

THE ELECTROSTATIC COMPONENT OF THE FORCE OF SLIDING FRICTION

By ROBERT SCHNURMANN, Dr. rer. nat., M.Sc.,
AND
ERIC WARLOW-DAVIES, B.A., D.Phil., B.Sc.
London Midland and Scottish Railway Company,
Research Laboratory, Derby

Communicated by M. G. Bennett ; *received* 1 *September* 1941

ABSTRACT. For jerky motion due to relaxation oscillation to occur, the {friction, velocity} characteristic of sliding must exhibit a non-linear falling section and a minimum. A similar type of characteristic belongs to any electrical discharge device with a practically infinite resistance below a definite breakdown potential at which it becomes converted into a conductor. It is suggested that, due to contact electrification, the electrostatic component of the force of sliding friction assumes appreciable proportions when the boundary layer has dielectric properties ; thus cycles of slow charging and rapid discharging occur, and cause relative motion to proceed by cycles of slow sliding and rapid slipping when the average velocity, determined by the rate of propulsion of one friction element, is smaller than the velocity for which the frictional force assumes the smallest value. Experiments are presented to demonstrate that jerking decreases when the dielectric breakdown strength of the lubricating material is diminished, and that with the same lubricant jerking decreases as the rate of propulsion is increased. In these experiments, solid metal cylinders of 5·33 cm. nominal diameter were forced through hollow steel cylinders of slightly smaller bore, with various lubricants, in a 15,000-kgm. wt. recording Buckton testing machine which was operated slowly and at a uniform rate. The electric discharge current which accompanied jerking, even when *both* friction elements were made of steel, was observed when the friction elements were insulated from the Buckton machine and connected to a ballistic galvanometer. In accordance with the electrostatic interpretation of jerky motion due to relaxation oscillation, the surfaces remained undamaged and Amontons' law was obeyed when heavy jerking took place, whereas the relative motion of the two friction elements proceeded smoothly when the surfaces became severely torn, and the tangential force increased at a rate more than proportional to the radial force.

---

§ 1. INTRODUCTION

It is well known that slow sliding can proceed by a series of jerks exhibiting the two characteristics of relaxation oscillations*—a fixed amplitude, and a period determined by a relaxation time for the building up of a limiting value. In the experiments described below, the amplitude of the jerks was observed to decrease when the dielectric breakdown strength of the material used as a lubricant was diminished, or when, with the same lubricant, the velocity of sliding was increased. Jerking was accompanied by an electric discharge through a galvanometer connected to the two friction elements. The expectation was confirmed that jerky motion was due to the characteristics of sliding and not to scoring. Amontons' law was satisfied, and the surfaces remained undamaged even when the tangential force of friction was a large fraction of the normal force, particularly when the jerks were very heavy. On the contrary, when severe tearing was observed, the relative motion between the two solid bodies had proceeded smoothly and did not obey Amontons' law. It is known from observations of contact electrification that an electrostatic force resisting sliding must exist if the boundary layer consists of dielectric or semiconducting matter. The above observations support the view that this electrostatic component may be an appreciable part of the force of sliding friction.

---

* A very readable account of relaxation oscillations has been published recently by G. F. Herrenden-Harker, 1940 (*Amer. J. Phys.* **8**, 1).

§ 2. THE FORCE OF FRICTION AS A FUNCTION OF THE VELOCITY OF SLIDING

Under certain conditions, the sliding motion of elastically restrained bodies does not proceed smoothly, but by jerks; the friction elements appear to " stick together ", while the elastic force increases but is balanced by the correspondingly increasing force of *static friction.* When the elastic force reaches the limiting value of the force of static friction, static friction ceases to balance the elastic force, and a rapid slip suddenly occurs (figure 1). In these cases, for small velocities, sliding friction is obviously smaller than the limiting value of static friction, and falls off with an increase in velocity. Haykin and his associates (1933, 1940) have pointed out that, for jerky motion due to relaxation oscillation

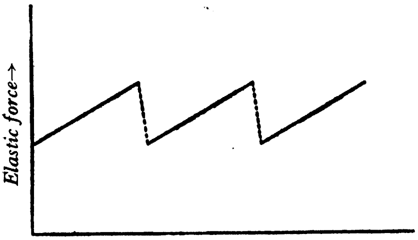


Figure 1

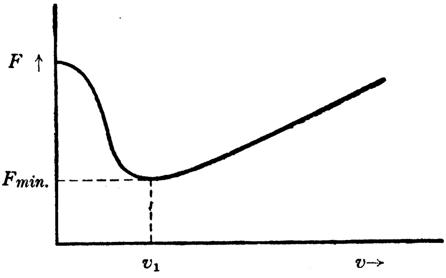


Figure 2

Figure 1. Force-time diagram when the average velocity of sliding is small and sliding friction is smaller than the limiting value of static friction. The two friction elements seem to stick together until the elastic force reaches the limiting value of static friction, when rapid slipping occurs.

Figure 2. Friction-velocity characteristic consistent with jerky sliding.

to occur, the {force of friction, velocity of sliding} characteristic of a particular frictional contact must exhibit the known characteristics of any mechanical or electrical device which is capable of performing relaxation oscillations under certain conditions, viz., a non-linear falling section and a minimum (figure 2). With this type of characteristic, two types of motion are possible, depending upon the uniform velocity v_0 at which one of the friction elements is moved forward. If this is given a value such that the representative point on the characteristic is to the left of the minimum ($v_0 < v_1$), the régime is unstable and oscillations occur. If the representative point lies on the rising arc to the right of the minimum ($v_0 > v_1$), the régime is stable and smooth sliding takes place.

Sliding proceeds by jerks when the régime is unstable. It begins when the elastic force reaches the limiting value of the force of static friction. On the falling section of the characteristic, the representative point upon reaching the unstable region flicks across almost instantaneously to the rising arc, which it follows towards smaller velocities, while the elastic force and the force of sliding friction both decrease. The representative point, upon reaching the unstable region near the minimum of the curve, again flicks almost instantaneously across to a small velocity (smaller than v_0), whence the elastic force increases until the limiting value of static friction is reached, and a new cycle begins. It is necessary to explain the given form of the dependence of frictional force upon sliding velocity, that is, the phenomenon of static friction and the falling section in sliding friction. This means that it is necessary to show which component of the force of friction has a characteristic with a non-linear falling section and a minimum, and to prove that under conditions of jerky motion this component represents an appreciable fraction of the frictional force.

§ 3. THE ELECTROSTATIC COMPONENT

Under certain conditions, electrical oscillations occur during sliding. It is well known that contact electrification arises whenever two different materials establish contact; when they are separated, they carry with them electrostatic charges of opposite sign (Schnurmann, 1941). When they are not separated, but made to slide one along the surface of the other, work must be done against the electrostatic attraction between the charges of opposite sign, if the surface layers have dielectric properties, so that the positions of the charges are fixed, and electrification follows processes of friction to the extent to which they involve making and breaking of contact. The separation of charges then increases as sliding proceeds until the potential reaches the breakdown value of the boundary layer, when this insulating layer suddenly becomes converted into a conductor, and a rapid discharge takes place. The {voltage, current} characteristic again has a non-linear falling section and a minimum, and (depending upon whether sliding is slow or rapid) the régime is either unstable or stable; that is, when sliding is slow it is accompanied by periods of slow charging and rapid discharge, whereas when sliding is rapid the discharge is continuous. Contact electrification during sliding is established, and the occurrence of discharges is known, for instance, when a glass rod is being rubbed with a woollen cloth; the experiments described below were designed to demonstrate that under certain conditions this electrostatic component bears an appreciable ratio to the force of sliding friction. The implication of the electrostatic interpretation of jerky motion—that the cycles of apparent stick and slip should be cycles of slow sliding and rapid slipping (slow charging followed by rapid discharge)—could not be tested to the extent of measuring the small amount of slow sliding, because in this region the internal displacements and forces mask the displacements and forces between the surfaces in contact when one of the friction elements is elastically restrained, and static methods are applied.

§4. EXPERIMENTAL METHOD

Metal pins were forced through steel rings with various lubricants in a 15×10^3-kgm. Buckton single-lever testing machine with balancing spring and gear for recording the force and the movement of the cross-head. The cylindrical surfaces of the friction elements were finished by grinding.

Various methods of " cleaning " the surfaces before the application of a lubricant were tried. Good reproducibility was obtained when the surfaces were rubbed with several pads of cotton wool which had been soaked in petroleum ether before each experiment, and when care was taken that the surfaces did not become oxidized while they were not in use. A coat of wool-fat protected them satisfactorily. The " cleaned " surfaces were either flushed with a liquid lubricant or wiped with a grease. Care was always taken that a surplus of lubricant surrounded the leading edge. In the case of volatile liquids, a dropping arrangement maintained a ring of liquid around the leading edge.

From the record of the maximum value of the force of friction and the calculated radial force, the coefficient of friction was computed. When the pin has fully entered the bore of the ring, the radial force is

$$P_r = \pi \times d \times p_r \int_0^{l_h} dl,$$

where d is the diameter of the bore, and for friction elements of the same material

$$p_r = \frac{E \times S}{d} \frac{(D^2 - d^2)}{2D^2}.$$

(E = Young's modulus, S = interference, D = outside diameter of the ring). The length l_h of the ground surface of the ring was always shorter than l_s, the length of the ground surface of the pin, so that the radial force increased linearly from zero at the beginning of assembling to a value which remained constant over a length $(l_s - l_h)$ and then decreased linearly to zero during dismantling. With steel elements the radial force in the fully assembled condition was, for instance, $P_r = 31{\cdot}4 \times 10^3$ kgm. wt. for $S = 2{\cdot}54 \times 10^{-3}$ cm., since the diameters were throughout $d = 5{\cdot}33$ cm. and $D = 10{\cdot}67$ cm. In the case of a brass pin and steel ring $P'_r = 25{\cdot}9 \times 10^3$ kgm. wt. would be calculated in the fully assembled condition from

$$p'_r = \frac{e}{\frac{1}{E_1}\left[\left(\frac{D^2+d^2}{D^2-d^2}\right)+\sigma_1\right]+\frac{1-\sigma}{E}},$$

where $e = S/d$, E_1 and E are Young's moduli of steel and brass respectively, and σ_1 and σ are Poisson's ratios for steel and brass. The pins were tapered at one end to ensure good entry conditions, and the rings were given a small radius at both ends.

The envelope of the recorded force of friction during a complete operation of assembling and dismantling of the friction elements at a uniform velocity of the Buckton machine (constant average rate of propulsion of one friction element) was represented by a trapezoid, if Amontons' law was obeyed. Deviations from Amontons' law were recorded as a more-than-proportionate increase of the force of friction with the path of sliding, i.e., the radial force.

In some series of experiments, the friction elements were electrically insulated from the Buckton machine by sheets of mica and were connected to a ballistic galvanometer * which indicated a discharge current with each jerk if the sequence of jerks was sufficiently slow for the mirror of the galvanometer to return to its zero position between two consecutive jerks. Otherwise the galvanometer would exhibit a constant deflection while jerky motion proceeded.

§ 5. RESULTS

Jerky motion and discharge currents. Characteristic results were obtained when a steel pin and ring were assembled with, for instance, butyl acetate, and when a similar pair of friction elements was coated with a machine oil, heated in a furnace with air circulation to 193° c. for, say, one to two hours and then assembled. In the former case, relative motion proceeded by a series of large jerks (figure 3), which increased with the nominal area of contact to reach a maximum value of 0·48 cm. in the fully entered condition. The envelope of the jerks is a trapezoid, the measured gradient being $2{\cdot}55 \times 10^3$ kgm. wt./cm. and the recorded maximum value of the force of friction $F_{\max} = 13{\cdot}6 \times 10^3$kgm. wt. Since $l_h = 5{\cdot}334$ cm., one would calculate $F_{\max}/l_h = 2{\cdot}55 \times 10^3$ kgm.wt./cm., in accordance with Amontons' law, which applies only to sliding that proceeds without tearing of the surfaces. Indeed, the dismantled surfaces did not show any sign of tearing in spite of the heavy jerks which had been observed throughout the process of assembling and dismantling of the elements. The behaviour of the other pair was quite different. No jerky motion was observed with it (figure 4); the force of friction increased at a larger rate than the nominal area of contact; the elements **seized**, i.e. the tangential force required to continue relative motion in the fully entered condition after a path of sliding of 6·3 cm. had been completed, was beyond the capacity of the machine. Inspection of the surfaces after the hollow element had been cut open, and thus separated from the solid one, showed that they had both suffered severe tearing* during a process of smooth continuous sliding. It is obvious from these two examples that tearing is not the result of jerky motion. On the contrary, a large number of experiments have shown that very heavy jerking without an appreciable deviation from Amontons' law is a certain indication that the surfaces are not severely torn during sliding at a small average value of the velocity.

---

* The authors are indebted to Mr. F. Record, of the Technical College of Derby, for the loan of a ballistic galvanometer.

When the pins and rings were electrically insulated from the Buckton machine and connected to a ballistic galvanometer, discharge currents were observed when jerking took place. In accordance with the view that the galvanometer deflections indicated the current which flowed when the contact potential reached the breakdown value of the boundary layer and ceased when the potential had fallen to a smaller value for which the resistance of the boundary layer was practically infinite, an appreciably larger galvanometer deflection was observed at the final jerk which made the pin drop out of the ring, because the complete separation of the two friction elements must lead to a *complete* discharge of their surfaces through the galvanometer to which they were connected. At a rate of propulsion of $2 \cdot 67 \times 10^{-3}$ cm./sec., the galvanometer deflection per jerk was observed to be 1·5 mm.† when a brass pin was assembled with a steel ring with Regelax medicinal paraffin-oil (Boots) as lubricant; the maximum value of the force of friction was $9 \cdot 5 \times 10^3$ kgm. wt., and the maximum jerk was 0·38 cm., both values being in close agreement with the results obtained with a steel pin and Regelax medicinal paraffin-oil as lubricant (see table 1). An increase in the rate of propulsion did not increase the galvanometer deflection, as was found when the Buckton machine was operated at a rate of $2 \cdot 67 \times 10^{-2}$ cm./sec., and the recorded maximum value

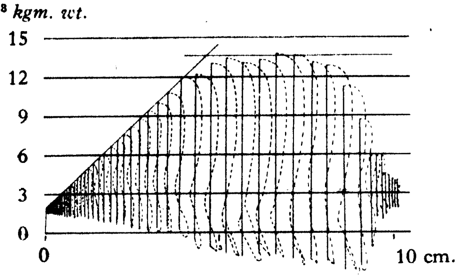


Figure 3.

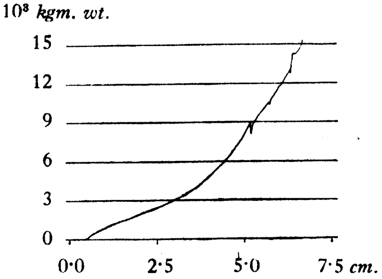


Figure 4.

Figure 3. Jerky motion when a steel pin and ring were assembled slowly and without tearing The frictional force (plotted on the ordinate) increased in proportion to the radial force (Amontons' law). The horizontal distance between the continuous almost-vertical lines indicates the magnitude of slipping. The dotted parts of the record are due to violent vibrations of the recording gear immediately after sudden slipping of the pin into the ring.

Figure 4. Smooth relative motion and deviation from Amontons' law when relative motion proceeded by severe tearing. The small kinks on the curve were obtained when the propelling operation was suspended.

---

* See *Engineering*, 1940, **150**, 236.

† This deflection was known from calibration to correspond to 4×10^9 coulombs.

of the force of friction of the solid element of brass dropped to $7\cdot82 \times 10^3$ kgm. wt. and the maximum jerk to 0·25 cm. At a still higher rate of propulsion, the discharge periods followed each other so frequently that the galvanometer assumed a steady deflection of, for instance, 2 mm. at $5\cdot64 \times 10^{-2}$ cm./sec., 3·5 mm. at $8\cdot75 \times 10^{-2}$ cm./sec., and 4·5 mm. at $1\cdot137 \times 10^{-1}$ cm./sec. In the last case a deflection by a further 11·5 mm. was observed at the instant of completed dismantling of the friction elements.

It is important to note also that a periodic discharge current was observed to accompany jerky motion when *both* the pin and ring were made of steel. In these particular experiments, Regelax medicinal paraffin-oil had been used as lubricant. The galvanometer deflection per jerk was 0·5 mm., i.e. smaller than with a brass pin, though of the same order of magnitude. This smaller value was probably due to the circumstance that these particular steel friction elements were defective, so that the maximum value of the force of friction was $7\cdot7 \times 10^3$ kgm. wt. instead of between $9\cdot8 \times 10^3$ kgm. wt. and $10\cdot2 \times 10^3$ kgm. wt., and the maximum jerk was 0·26 cm. instead of 0·41 cm. with undamaged steel elements (table 1). These " defective " steel elements had been used previously eight times, and had acquired one bad score during the eighth series of experiments. A shortage of specimens made it necessary to recondition this pair by filing over the scored portions of the surfaces and to use them in the " defective " condition. Thus, since both the friction and the jerks were smaller than with undamaged surfaces, the galvanometer deflection per jerk would also be expected to be smaller. Anyhow, the essential point is that such deflections were clearly observed when both the pin and ring were made of steel.

Further evidence of the galvanometer deflection as an indication of a discharge current was obtained at the high rates of propulsion when the magnitude of the steady deflection varied with the nominal area of contact between the friction elements. For the reason mentioned above (that the contact potential would vary in the assembled condition between the breakdown value and a slightly smaller value for which the resistance of the boundary layer was practically infinite, and between the extinction value and zero when separation of the surfaces took place), the steady deflections observed at the high rates of propulsion were largest during the stage of dismantling and smallest in the fully entered condition. At a rate of propulsion of $8\cdot75 \times 10^{-2}$ cm./sec., for instance, a deflection of 3 mm. was observed in the assembling stage, a deflection of 2·5 mm. in the fully entered condition, and a deflection of 3·5 mm. during dismantling. The corresponding observations at a rate of propulsion of $1\cdot137 \times 10^{-1}$ cm./sec. were 4 mm., 3 mm., and 4·5 mm. respectively.

Variation of the dielectric breakdown strength of the lubricant

A brass pin and a steel ring were assembled once with Regelax medicinal paraffin-oil, and another time with a mixture of 50 % rape oil and 50 % Russian tallow. The rate of propulsion was $2\cdot67 \times 10^{-3}$ cm./sec. in both cases. In the

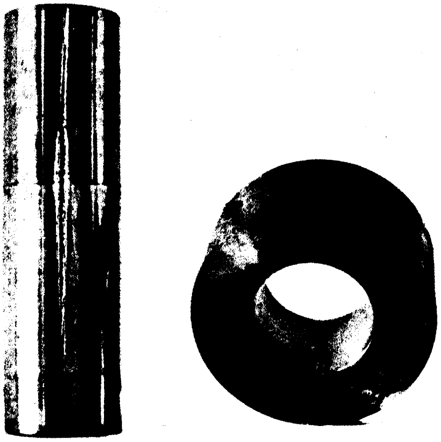


Figure 5. Damage to the surfaces of the pin and ring after smooth relative motion. The assembly was made with pure caprylic acid which severely attacked the steel, as the top surface of the ring shows.

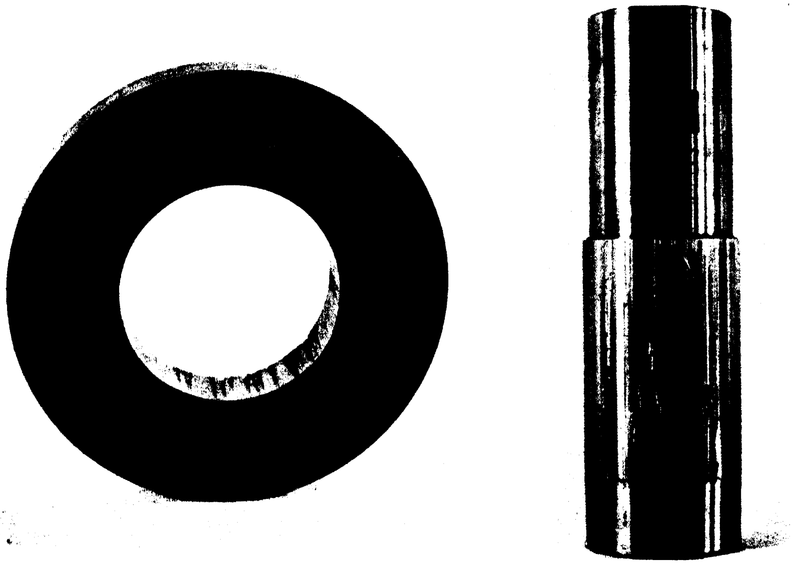


Figure 6. Severe tearing of the pin and ring after assembly with a mixture of two parts of Regelax medicinal paraffin-oil with one part of octyl alcohol, which attacked the steel, as the top surface of the ring indicates. The blobs of metal which can be seen on the surface of the pin were not welded on and could easily be removed.

Table 1

Lubricant	Voltage		Interference (10^{-3} cm.)	Rate of propulsion (10^{-3} cm./sec.)	$F_{\max}$ (10^3 kgm. wt) (observed)	$\mu_{\max} = F_{\max}/P_r$ (computed)	Max. jerk (cm.) (observed)
	Held for 1 min. (kilovolts)	Breakdown (kilovolts)					
Regelax	40	45 (instantaneous)	2·5	2·54	10·1	0·32	0·4
,,	40	45 (instantaneous)	2·5	2·54	9·8	0·31	0·42
,,	40	45 (instantaneous)	2·5	2·54	10·2	0·32	0·41
Castrol XL	50	50 (almost instantaneous)	2·5	2·54	8·7	0·28	0·38
Soft soap	—	—	2·5	2·54	7·5	0·24	0·25
Regelax + 1 % caprylic acid	15	16 (after 15 sec.)	2·5	2·54	6·5	0·21	0·18
Regelax + 5 % technical oleic acic	—	—	2·8	2·67	6·51	0·19	0·15
Castor oil	—	—	2·8	2·67	6·18	0·18	0·20
Raw linseed oil	—	—	2·8	2·67	5·62	0·16	0·05
Rape oil	30	35 (after 40 sec.)	2·8	2·67	5·21	0·15	0·06
Tallow and rape oil	25	30 (after 17 sec.)	2·5	2·67	5·43	0·17	0·14

latter case, both the force of friction and the magnitude of the jerks were smaller, and the galvanometer deflection per jerk was 0·25 mm. instead of 1·5 mm. with Regelax. At a rate of propulsion of $1{\cdot}137 \times 10^{-1}$ cm./sec., the galvanometer showed a steady deflection of 4 mm. when the mixture of tallow and rape oil was used as lubricant, and a large increase of this deflection was observed when the two friction elements parted contact.

The friction results obtained with steel elements and various lubricants were compared with the dielectric breakdown strengths of these materials in bulk (table 1).

Very similar friction results had been obtained with proprietary samples of Regelax medicinal paraffin-oil and Castrol XL motor oil (Wakefield) before their dielectric breakdown strengths were measured. A different sample of Castrol XL which had been stored in the works was sent to the chemical laboratory,* where it was found that 11 kilovolts were held for one minute (transient sparking) and that instantaneous breakdown occurred at 15 kilovolts. The chemists, who did not know of the friction results, remarked in their report that the presence of a trace of water in this oil was suspected. They were then supplied with a proprietary sample of Castrol XL and found 50 kilovolts, a result very close to the one obtained with Regelax, and in agreement with the observation that both these lubricants had given similar friction results.

Table 2

$F_{\max}$ (10^3 kgm. wt.) (observed)		$\mu_{\max} = F_{\max}/P_r$ (computed)		Max. jerk (cm.) (observed)	
Castrol High-Press		Castrol High-Press		Castrol High-Press	
Unfiltered	Four times filtered	Unfiltered	Four times filtered	Unfiltered	Four times filtered
5·10		0·15		0·08	
	6·27		0·18		0·175
	5·7		$0{\cdot}1_7$		0·09
5·7		$0{\cdot}1_7$		0·07	
6·0		$0{\cdot}1_7$		0·09	
	6·56		0·19		0·12

The dielectric properties of oils can as a rule be slightly improved by filtration, and the claim has been made recently by du Noüy (1940) that the lubricating properties of, for instance, a good motor oil are diminished when it has been filtered four times through paper. Experiments with Castrol High-Press lubricating oil (Wakefield), both in the " as received " condition and after four filtrations through paper showed larger values both of the force of friction and of the jerks in the latter case (table 2). All but the first of the results summarized in table 2 were obtained with the same pin and ring; the interference was $2{\cdot}8 \times 10^{-3}$ cm. throughout.

---

* The authors express their thanks to Mr. E. A. Coakill for the measurement of the dielectric breakdown strengths in accordance with the B.S.I. specification for Insulating Oils for Electrical Purposes (No. 148, 1933).

Sir William Hardy, as is well known, deduced from experiments with various members of a chemical series among the substances used as lubricants that the friction decreased with increasing molecular weight in any given series. This relationship was found to hold for series of acids, alcohols and paraffins. For substances of the same molecular weight but in different series the friction was found to be dependent upon the chemical structure of the molecule. Fogg (1940) measured static friction with a modified form of Deeley machine, using as lubricants methyl and ethyl esters which by virtue of their chemical inactivity were suitable as regards freedom from surface corrosion. He noticed a maximum value of the coefficient of static friction at molecular weights of about 88 and 120 respectively, and suggested an explanation of this phenomenon by making certain assumptions regarding the structure and method of attachment of the molecules to the surface. In the present experiments with pins and rings both made of steel (interference: $2{\cdot}8 \times 10^{-3}$ cm.), and methyl and ethyl esters as lubricants, at a rate of propulsion of $2{\cdot}67 \times 10^{-3}$ cm./sec., the maximum values of the force of friction were found to decrease in both series with increasing molecular weight (table 3), but when esters with the same acid radical and various alcohol radicals were compared (table 4), the rule did not apply that the friction decreased with increasing molecular weight.

Table 3

Lubricant	Molecular weight	$F_{\max}$ (10^3 kgm. wt.) (observed)	$\mu_{\max} = F_{\max}/P_r$ (computed)	Max. jerk (cm.) (observed)
Methyl formate	60	13·0	0·38	—
,, acetate	74	11·8	0·34	0·09
,, propionate	88	8·9	0·26	0·25
,, butyrate	102	8·0	0·23	0·18
,, valerianate	116	8·7	0·25	$0{\cdot}22_5$
Ethyl formate	74	13·3	$0{\cdot}3_9$	0·3
,, acetate	88	12·6	$0{\cdot}3_7$	0·45
,, propionate	102	9·83	0·28	$0{\cdot}32_5$
,, lactate	118	9·54	0·28	$0{\cdot}27_5$
,, iso-valerate	130	9·17	$0{\cdot}2_7$	$0{\cdot}27_5$
,, oenanthate	158	7·8	0·23	0·10
,, capryllate	172	6·17	0·18	0·10
,, stearate	312	5·7	0·17	$0{\cdot}07_5$

The electrical conductivities at 25° c. of methyl acetate ($\kappa = 34 \times 10^{-7}$) and of ethyl acetate ($\kappa < 1 \times 10^{-9}$) are recorded in the literature (*International Critical Tables*, 1929, vol. 6, p. 143). If the dielectric breakdown strength can be assumed to be inversely proportional to the conductivities, the friction ought to be larger with ethyl acetate as lubricant in accordance with the experimental results (table 3). From some relative values of the electrical conductivities of ethyl esters at room temperature (*op. cit.* p. 145, figure 2) one might expect the friction to be larger with ethyl acetate as lubricant than with ethyl propionate (table 3). However, the comparison of all the experimental results obtained with esters with the recorded relative values of the electrical conductivities would also suggest some inconsistencies. The possibility was not ruled out that the

Table 4

Lubricant	Molecular weight	F_{max} (10^3 kgm. wt.) (observed)	$\mu_{\text{max}} = F_{\text{max}}/P_r$ (computed)	Max. jerk (cm.) (observed)
Methyl acetate	74	11·8	0·34	0·09
Ethyl ,,	88	12·6	$0{\cdot}3_7$	0·45
Propyl ,,	102	11·2	0·32	0·36
Butyl ,,	116	13·6	0·39	0·48
Iso-propyl ,,	118	13·3	$0{\cdot}3_9$	0·46
Amyl ,,	130	12·0	0·35	0·4
Propyl phthalate	250	11·2	0·32	0·36
Butyl ,,	278	11·3	0·33	0·45
Amyl ,,	306	11·6	0·34	$0{\cdot}42_5$

Table 5

Lubricant	Interference (cm.)	Rate of propulsion (10^{-3} cm./sec.)	F_{max} (10^3 kgm. wt.) (observed)	$\mu_{\text{max}} = F_{\text{max}}/P_r$ (computed)	Max. jerk (cm.) (observed)
Machine oil	$2{\cdot}03\times10^{-3}$	2·54	5	0·2	0·2
		5·08	—	—	0·2
		10·2	4	0·16	$0{\cdot}12_5$
		31·7	2·5	0·1	$0{\cdot}02_5$
Machine oil (heated to 200° c. for 30 min. while air bubbled through)	$2{\cdot}54\times10^{-3}$	2·54	5	0·16	0·12
		5·08	4·7	0·15	0·1
		10·2	4·38	0·14	0·1
		31·7	3	0·1	
Regelax	$2{\cdot}54\times10^{-3}$*	2·67	9·54	0·37	0·35
		26·7	7·82	0·30	0·25
		56·4	6·32	0·24	0·17
		87·5	6·0	0·23	$0{\cdot}17_5$
		113·7	5·16	0·2	$0{\cdot}07_5$
Regelax	$7{\cdot}62\times10^{-3}$†	2·54	5·5	$0{\cdot}1_8$	0·16
		5·08	5	$0{\cdot}1_6$	0·09
		10·2	4·54	0·14	0·07

*** The solid element was made of brass.**

† The length of the ground surface of the ring was 1·778 cm. The ring became stretched in the course of the experiment.

volatility of the lower members of the series might be responsible for these. It will be recalled that the ground surfaces of a pin and ring were flushed with an ester immediately before they were assembled, and that in the case of volatile liquids a dropping arrangement maintained a surplus of lubricant around the leading edge. The possibility that in these cases the ground surface of the ring dried up while the pin entered cannot be excluded.

Variation of the rate of energy supply

It was mentioned above that the rate of energy supply to a device with a characteristic exhibiting a non-linear falling section and a minimum would decide whether the functioning of the device was unstable or stable. In the present case, the velocity at which the Buckton machine is operated, i.e. the rate of propulsion of one friction element, represents the rate of energy supply. As this increases from zero, the oscillations, i.e. the jerks, must decrease in the unstable region, and must give way to continuous motion (smooth sliding) at a sufficiently high velocity for stability to be ensured. Table 5 illustrates the decrease of both the maximum value of the force of friction and the jerks with increasing rate of propulsion.

Oxidation and tearing

The electrostatic component of the force of sliding friction decreases as the dielectric breakdown strength of the boundary film is reduced, and the jerks become correspondingly smaller. This is the case, for instance, when one per cent of caprylic acid is added to Regelax medicinal paraffin-oil and sliding proceeds without appreciable damage to the metal surfaces, so that Amontons' law is obeyed. No jerks were observed when pure caprylic acid was used as lubricant for friction elements of steel. But in this case the metal surfaces (interference: $2 \cdot 54 \times 10^{-3}$ cm.) suffered severe tearing (figure 5), and the force of friction reached 7×10^3 kgm. wt. (rate of propulsion: $2 \cdot 67 \times 10^{-2}$ cm./sec.) in the fully entered condition of the solid element. The caprylic acid was found to attack the steel. Tearing and deviations from Amontons' law were always observed when either the material used as lubricant appreciably attacked the metal surfaces or when these had become discoloured with corrosion before the lubricant was applied. For instance, steel friction elements which had been standing without a coat of wool-fat after the grinding operation had become slightly discoloured with corrosion before they were assembled with tallow and rape oil. Very few small jerks occurred, and the surfaces suffered severe tearing. Standard soap solution, and a mixture of two parts of Regelax medicinal paraffin-oil with one part of octyl alcohol when used as lubricants also gave rise to severe tearing (see figure 6). The friction records exhibited appreciable deviations from Amontons' law. The plucked out particles were easily levered off the surfaces; they were *not* blobs of welded-on metal.

§6. THE LIMITING VALUE OF THE FORCE OF "STATIC FRICTION"

As a result of static friction, the elastic force increases as one friction element is moved forward at a uniform velocity while the other one is elastically restrained. When the elastic force reaches the limiting value of the force of static friction, the slider must begin to move, because for small velocities of sliding the force of friction is smaller than the limiting value of static friction, and falls off with an increase in velocity in those cases in which jerky motion is observed. The electrostatic interpretation of jerking would require that the cycles of sticking and slipping should be, in fact, cycles of slow charging and rapid discharging, i.e. cycles of slow sliding and rapid slipping. This would mean that the limiting value of static friction was reached by elastic deformation, together with a slight actual displacement at the contact area of the two bodies, which would be accompanied by increased charge separation until the contact potential reached the breakdown value. The slight displacement preceding slipping was too small to be noticed with the sensitivity of the measuring equipment fitted to the testing machine. The elastic distortion of the machine and the specimens was measured in a separate experiment, when a steel plate prevented the pin from entering the ring, with the result that it amounted to $7{\cdot}5 \times 10^{-3}$ cm. for every 1×10^3 kgm. wt.

In addition to the above-described experiments, which support the view that the electrostatic component of the force of sliding friction can assume appreciable proportions if the charge separation increases as sliding proceeds, the contribution of the electrostatic force to the radial force can also be estimated. Helmholtz assumed that when two dissimilar surfaces are in contact, an electrical double layer of about 10^{-8} cm. in thickness exists between them. The difference between the breakdown potential and the truly static value might be assumed to be one-tenth of a volt, so that a radial electrostatic force of 4×10^3 kgm. wt. in the fully entered condition would be computed from

$$P_{el} = \frac{\pi \times d \times l_h}{8\pi(10^{-8})^2} \, \frac{(10^{-1})^2}{8{\cdot}83 \times 10^{10}}.$$

If the electrostatic interpretation of the friction characteristic with a non-linear falling section and a minimum is correct, no jerking can be expected when two *naked* metal surfaces are in frictional contact, because in this particular case the charge separation does not increase as sliding proceeds, since the positions of the charges are *not* fixed in the metal surfaces. It is a matter of considerable experimental difficulty to work with naked metal surfaces, but under conditions of severe tearing the re-contamination of the freshly created contact area from the surrounding atmosphere may require a short time. Bowden and Leben (1939) reported that an elastically fixed platinum slider in frictional contact with silver in the air of the laboratory exhibited slight jerking at the beginning of the run (the friction element of silver was moved forward at a uniform velocity of a few thousandths of a centimetre per second). "The friction was low", the coefficient being about 0·6 (Bowden, Leben and Tabor, 1939), and a smooth groove was cut in the silver. After relative motion had proceeded over a short distance, the behaviour of the motion changed, jerking disappeared, and " the friction rose to a high value ", the coefficient fluctuating now between about $1{\cdot}2_5$ and $1{\cdot}3_5$. A close examination showed that a small blob of silver had become attached to the platinum surface, and it would appear that the friction characteristic of the naked metal-to-metal contact created by plucking the blob of silver from the silver plate did not allow of jerking, until the inevitable re-contamination by matter adsorbed and condensed from the air of the laboratory had been completed.

§ 7. DISCUSSION

A different interpretation of sliding which proceeded by a series of jerks was proposed by Bowden and Leben, who suggested that the friction behaviour should depend upon the relative physical properties of the metals, particularly on the melting point rather than upon the properties of the boundary layer. Friction should be due to a welding together of the metals at the local points of contact, and the metallic junctions, when broken, should distort the metal " to a considerable depth ". In their detailed discussion of Bowden and Leben's suggestion, Haykin and his collaborators pointed out that local melting—if it does take place—would be the *result*, and not the *cause* of jerky motion. But even as the *result* of short-lived slipping, no large temperature rise can be expected, and it is well to bear in mind that at every stage the temperature of the surface irregularities is determined by the equilibrium between the amount of heat generated and that dissipated, and that the dissipation of heat by conduction, radiation and convection also increases with the temperature, so that, as Bowden and Ridler (1936) have shown experimentally, *large* values both of the force of friction and of the speed of sliding are required to achieve an appreciable temperature rise (Schnurmann, 1939). The temperature rise under the conditions of Bowden and Leben's experiments can be estimated to be about 1° c. if each jerk is completed in one-thousandth of a second. For this estimate, the results obtained by these authors with silver sliding on a steel plate which was moved forward at a velocity of 6×10^{-3} cm./sec. were used. The silver slider was assumed to be a sphere of 0·3 cm. diameter, and the area of contact between silver and steel was estimated from the Brinell hardness of silver.

The present experiments show clearly that jerking is due to the characteristics of sliding, and not to scoring. When the conditions are such that the electrostatic component of the force of sliding friction is appreciable, and varies as sliding proceeds, jerky motion must arise. This explains Kaidanovsky and Haykin's observation that the friction characteristic of a pair of friction elements of wood and a metal has a similar form to that of two dissimilar metals in the air of the laboratory.

One experiment was made with the intention of reducing the jerks by passing an electric current through the frictional contact. $F_{\max} = 8{\cdot}5 \times 10^3$ kgm. wt. was found when a steel pin and ring with an interference of $2{\cdot}8 \times 10^{-3}$ cm. were assembled with Regelax medicinal paraffin-oil at a rate of propulsion of $2{\cdot}67 \times 10^{-3}$ cm./sec., and currents up to 2 amp. passed the contact. The maximum jerk was 0·23 cm. instead of 0·41 cm. (table 1) without the current. Much higher current densities would be required for obtaining smooth sliding due to dielectric breakdown of the entire boundary layer.

§ 8. ACKNOWLEDGEMENT

The authors express their thanks to the London Midland and Scottish Railway Company for permission to publish this paper.

REFERENCES

Bowden, F. P. and Leben, L., 1939. *Proc. Roy. Soc.* A, **169**, 371.
Bowden, F. P., Leben, L. and Tabor, D., 1939. *Engineer*, **168**, 214.
Bowden, F. P. and Ridler, K. E. W., 1936. *Proc. Roy. Soc.* A, **154**, 640.
Fogg, A., 1940. *Proc. Phys. Soc.* **52**, 239.
Haykin, S., Lissovsky, L. and Solomonovich, A., 1940. *J. Phys. U.S.S.R.* **2**, 253.
Kaidanovsky, N. L. and Haykin, S., 1933. *J. Techn. Phys.* **3**, 91.
du Noüy, P. L., 1940. *C.R. Acad. Sci., Paris*, **209**, 101.
Schnurmann, R., 1939. *Engineer*, **168**, 278.
Schnurmann, R., 1941. *Proc. Phys. Soc.* **53**, 547.

THE ABSOLUTE MEASUREMENT OF ELECTRICAL RESISTANCE BY A METHOD USING THE AVERAGE ELECTROMOTIVE FORCE OF A COMMUTATING GENERATOR*

By H. R. NETTLETON, D.Sc.
AND
E. G. BALLS, M.C., B.Sc., Ph.D., A.I.C.,
Birkbeck College

Received 25 *September* 1941

ABSTRACT. An electrical resistance of approximately 1 ohm is measured absolutely with rapidity and precision in terms of a mutual inductance and a frequency, by balancing the e.m.f. across it when conveying direct electric current, against the average e.m.f. of a commutating generator, the field coils of which are traversed by the same current.

By a new method, a remarkable "flat" of mutual inductance is obtained between the rotor and the field coils over a range of 20° of arc in the neighbourhood of the commutations. Further, over an adjustable portion of this range, covering the actual breaks at commutations, the rotor together with additional adjustable resistance is short-circuited,

---

[*] A preliminary investigation of this method has been described by one of us (E. G. B.) in Part II of a Thesis for the Ph.D. degree of the University of London.

thus enabling the e.m.f. across the resistor to suffer no interruption in its supply to the detector and allowing it to supply the extra quantity of electricity absorbed by the rotor, in virtue of its self-inductance, on establishing the current through it at each commutation. In this way the two main errors associated with break and self-inductance are set against one another, and compensation is readily effected by the adjustment of a resistance in a preliminary experimental test carried out under conditions in which these two effects preponderate.

The rotor has a constant speed of 12·5 revolutions per second, and the maximum mutual inductance of some 20,000 μH. permits of rapid and fine adjustment over a range of 400 μH. The balance is sensitive to 0·1 μH. and an accuracy of a few parts in a hundred thousand has been obtained.

---

§ 1. INTRODUCTION

THE method of measuring an electrical resistance absolutely by balancing the e.m.f. across it when conveying direct electric current, against the average e.m.f. across a commutating generator, the field coils of which are traversed by the same current, was suggested by Rosa (1909). Although Rosa proposed a form of apparatus and a method of procedure, difficulties were encountered and no experimental data, derived from any form of this method, have hitherto been published (Curtis, 1937).

The method is attractive on account of the high sensitivity, due to the relatively high e.m.f. involved, and because of the simplicity of the basic expression for R, the resistor to be measured, viz.,

$$R=4n \,.\, M_0, \qquad \ldots\ldots(1)$$

where n is the frequency of revolution of the generator and M_0 is the maximum mutual inductance between the rotor and the field coils.

On the other hand, before the method can attain precision value, three sources of error associated with the break at commutation must be overcome. The first of these, due to the interruption of the e.m.f. of the generator, is adequately rendered negligible by arranging that the mutual inductance between the rotor and the field coils is very flat around the maximum value, so that over the range of commutation the e.m.f. is sensibly zero. The remarkable nature of the "flat", described in this paper, and the additional safeguard associated with the short-circuit device and the method of measuring the effective mutual inductance do away with this trouble completely. A much more serious error, however, may be expected to arise from the interruption through the detector of the steady e.m.f. drawn off the resistor under test, while a third source of error which might be considerable is associated with the effect of self-inductance of the rotor, through which a current is re-established at each commutation.

In order to overcome the major difficulties, Rosa suggested the use of a differential galvanometer as the detector, one coil of which is in uninterrupted connection with the resistor, while other coils are connected suitably with his proposed two-phase composite rotor. This use of a differential galvanometer, however, which permits quantities of electricity, whose aggregate values over a period of revolution are large, to flow through the various coils, adds considerable complication to the method and prevents the realization of a simple null balance independent of the resistance of the detector circuit.

In the method here described, the e.m.f. across the resistor is directly balanced for zero aggregate quantity of electricity through the galvanometer (over a period of revolution and very approximately over half a period) against the average e.m.f. of the commutating generator. By a new and simple method, the rotor is made to have an extraordinarily flat maximum of mutual inductance with the field coils over some 20° of arc in the neighbourhood of commutation. Further, over an adjustable portion of this range of some 5° to 8°, completely covering the actual commutations, the rotor, together with some small resistance, is short-circuited out of the detector loop, thus enabling the e.m.f. across the resistor to function through the detector without interruption and to supply the extra quantity of electricity absorbed by the rotor at each commutation on account of its self-inductance. In this way, the two main errors arising from the break of the resistor and from the self-inductance of the rotor may be set against one another and compensation is readily effected by experimental test under conditions in which these effects are magnified.

A form of apparatus has been designed for the specially accurate measurement of resistances of approximately 1 ohm, thus enabling the international ohm to be measured in C.G.S. units. For this purpose, the rotor is driven by a synchronized motor at a constant speed of 12·5 rev. per sec. which may be checked by a 1000-cycle note, and the mutual inductance of some 20,000 μH. between the rotor and the field coils at the boundaries of short circuit is varied for final balance on a high-sensitivity galvanometer. The balance, sensitive to 0·1 μH., is obtained and confirmed in a few minutes; the rotor is then stopped and set by an electrical contact test at each in turn of the four limiting positions of short-circuit, where the mutual inductance is measured to 0·1 μH. under the same direct primary current as that used throughout the main experiment, by balance against a 10 mH. certified standard together with a trifle over 10 mH. on a Campbell mutual inductometer which can readily be checked against the standard. The sum of the four readings represents the flux change per unit current per revolution and replaces $4M_0$ in the basic equation (1), which is shown below to be accurate under the experimental conditions to about three parts in a million.

§ 2. DESIGN OF A SUITABLE GENERATOR

Attainment of a flat maximum of mutual inductance between a rotor and twin field coils

Rosa (1909) proposed the use of twin field coils set somewhat further apart than the Helmholtz distance of separation, with a rotor of large diameter set symmetrically between them. A more convenient and efficient arrangement consists of two twin field coils (A and B, figure 1) separated by any small distance and two relatively small twin-rotor coils, C and D, with their planes parallel and displaced by a distance *d* from the axis of rotation O. The distance *d* is approximately equal to half the radius of the field coils, and by slight adjustment of this distance the maximum "flat" of mutual inductance is readily obtained. The actual rotor is seen in figure 8. The degree of "flat" of mutual inductance which may be obtained around the maximum is illustrated by the curves of figure 2. These were obtained with our twin field coils of 16·4 cm. mean radius,

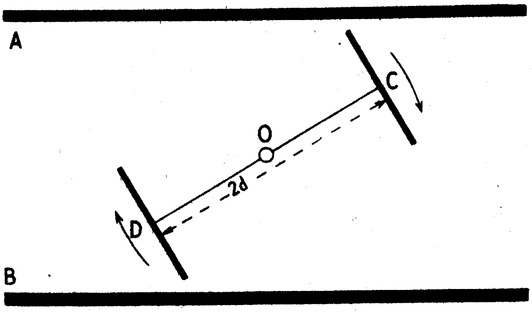


Figure 1. Form of generator.

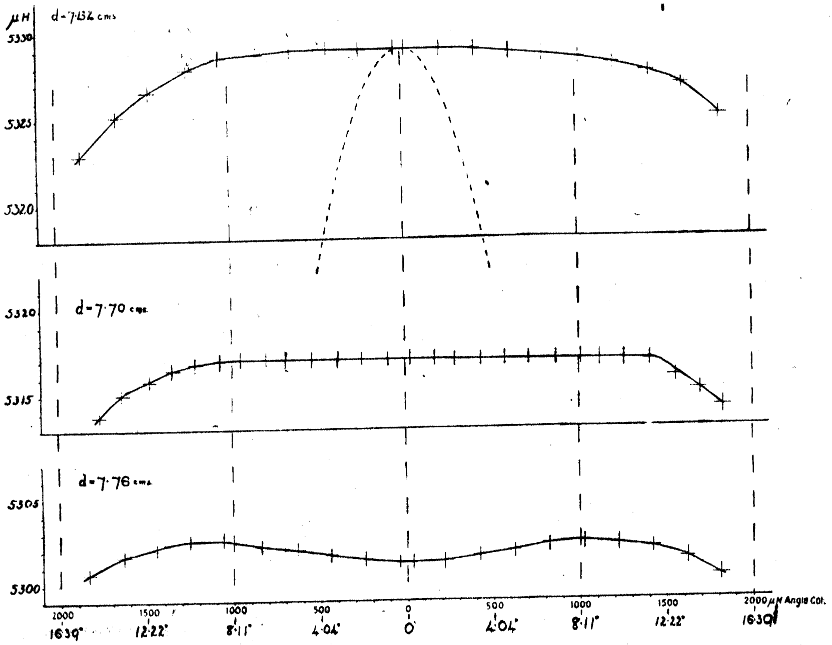


Figure 2. The "flat" of maximum mutual inductance between a rotor coil and field coils for three different displacements of the rotor coil from the axis of rotation. By contrast, the sine law is shown by the dotted curve.

and one of the coils C of the rotor used in the preliminary investigation. This coil, of 334 turns of d.s.c. copper wire of S.W.G. 26 was wound on a former having a channel 1·2 cm. wide and an internal diameter of 7·6 cm. These curves show the variation with angular displacement of the mutual inductance between this rotor coil and the field coils at three slightly different distances *d* of displacement of C from the axis of rotation. The contrast from the sine law, which holds approximately when *d* is zero, is shown for the same maximum by the dotted curve. In the position of maximum flat, the extreme variation of mutual inductance is only 0·2 μH. in 5317 μH. over a range of 20° of arc, and a setting of the rotor coil between this position and that for the lowest curve is ideal. The coil D is similarly set and then joined in series and in conjunction with C. The angular displacements, though not required accurately, were most conveniently measured by an angle coil necessary for the wave-form determinations below (Llewellyn, 1934; Nettleton and Balls, 1935; Balls, 1938).

Wave form of the e.m.f. of the generator and of the currents through the detector and rotor at null balance

With the aid of an angle coil obeying the sine law, inserted in the hole h of the oak bed (figure 8) of the rotor used in the preliminary investigation, the mutual inductance M between the rotor and the field coils was explored at various displacements θ from the conjugate position. The variation of M with θ enabled the corresponding values of $dM/d\theta$, proportional to the generated e.m.f. under constant speed of revolution, to be deduced and plotted against θ as a wave form.

In the actual experimental measurement of resistance, a balance for zero aggregate quantity of electricity through the detector must be obtained from the two opposing e.m.f.s. shown in figure 3, the one, a steady e.m.f. drawn off the resistor under test (ignoring the effects of break and self-inductance), and the other, the variable e.m.f. of the generator, having zero value in the neighbourhood of commutation.

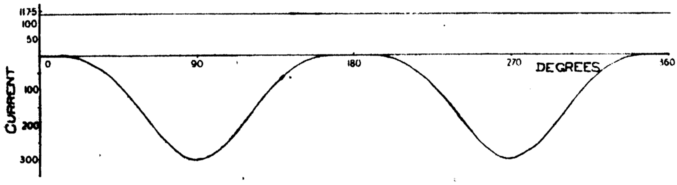


Figure 3. The steady e.m.f. opposed by a variable e.m.f. from the generator.

The resultant wave form of the current through the detector at balance is shown in figure 4, and has a period equal to half the time of revolution of the rotor; it is important in the complete theory of the method given below. Over a range of about 50°·5 on each side of the maximum of mutual inductance, a current flows through the detector in a positive direction, the flat maximum value, c_0, being equal to that through the detector when the rotor is at rest; over the range of 39°·5 on each side of the zero of mutual inductance and maximum e.m.f. of the generator, an equal quantity of electricity flows through the detector in the negative direction.

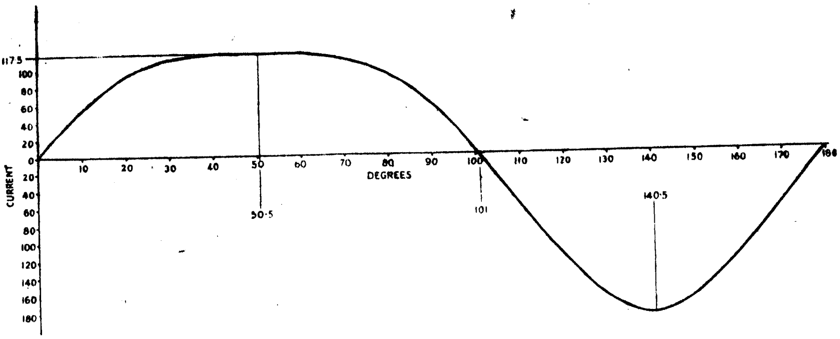


Figure 4. The wave form of the current through the detector at balance.

The wave form of the current through the rotor at balance for zero aggregate quantity of electricity through the detector is shown in figure 5, and has a period equal to the time of revolution of the rotor. The current c_0 is reversed through the rotor at each commutation.

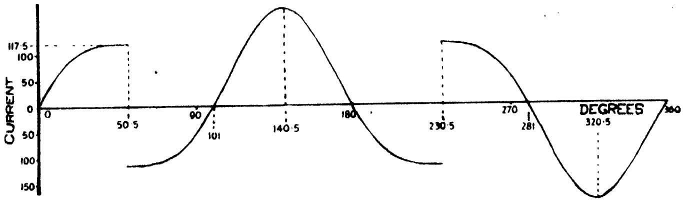


Figure 5. The wave form, at balance, of the current through the rotor.

The effects of the short circuit covering the commutations and of self-inductance are considered in the next section.

§ 3. THEORY OF THE METHOD

The circuit and fundamental equations

In the arrangement shown in figure 6, the primary circuit consists of a steady voltage E in series with the resistor R to be measured and additional resistance X, which includes the field coils of the generator and rheostats of total self-inductance μ. The secondary circuit drawn off the potential leads of R is of total resistance $R+F$. The resistance F is of at least 8000 ohms, and includes a high-resistance galvanometer G of self-inductance λ, additional non-inductive resistance, and the resistance r, which comprises the rotor of resistance

$r_1 = 17{\cdot}5$ ohms and self-inductance $L = 0{\cdot}0385$ H. and the additional non-inductance resistance r_2 in ohms and tenths up to 20 ohms. The shorting rotor N is adjustable and enables the commutations to be covered over various angles of from 4° to 9° of arc; when operative, it cuts out a resistance $r = r_1 + r_2$ from the secondary circuit.

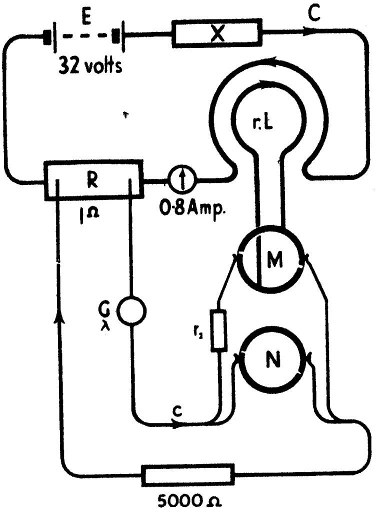


Figure 6. The essential circuit.

Let C and c respectively denote the instantaneous values of the currents in the primary and secondary circuits, and M the instantaneous mutual inductance between them. Then, with the short-circuit inoperative, we have, by applying Kirchhoff's laws to the circuits,

$$CX + (C - c)R = E - d(cM)/dt - \mu\,.\,dC/dt, \qquad \ldots\ldots(2)$$

$$cF - (C - c)R = -d(CM)/dt - (\lambda + L)dc/dt, \qquad \ldots\ldots(3)$$

whence

$$\begin{aligned} c(FX + FR + XR) = ER - (X + R)\,.\,d(CM)/dt \\ - (X + R)(\lambda + L)dc/dt - Rd(cM)/dt - R\mu dC/dt, \qquad \ldots\ldots(4) \end{aligned}$$

which may be written in very convenient form, showing the quantity of electricity dQ traversing the detector in the time dt, viz.,

$$dQ = c\,.\,dt = [ER\,.\,dt - (X + R)\,.\,d\phi - R\,.\,d\psi]/Y, \qquad \ldots\ldots(5)$$

where $d\phi$ = flux change in the detector circuit in the time dt,
$d\psi$ = flux change in the primary circuit in the time dt,

and $$Y = FX + FR + XR. \qquad \ldots\ldots(6)$$

Owing to the large " flat " of maximum mutual inductance, no flux changes occur over several degrees surrounding the short circuit and commutation, and

C, c and M are constant. The limiting values of C and c just before and just after short circuit are then given by

$$C_0 = E(F+R)/Y, \qquad \ldots\ldots(7)$$

$$c_0 = ER/Y. \qquad \ldots\ldots(8)$$

Elsewhere, the value of C only differs minutely from the value C_0, but the variation of c at balance is that shown by the wave form of figure 4.

There are various flux changes and transients associated with the commutations and short circuits that cover them, but the most evident effect is the modification of the current c_0 due to the removal of the resistance r from the secondary circuit. On short circuit, the current through the detector, almost instantaneously attains a value c_s given by

$$c_s = ER/Y_s, \qquad \ldots\ldots\ldots(9)$$

where

$$Y_s = (F-r)(X+R)+XR; \qquad \ldots\ldots(10)$$

and hence, using (6) and (8),

$$c_s = c_0\left[1+\frac{r}{F-r+RX/(X+R)}\right]. \qquad \ldots\ldots(11)$$

Now $F+RX/(X+R)$ is the resistance of the detector circuit with the rotor unshunted and with the primary circuit closed as a shunt across R; whence, writing F_0 for $F+RX/(X+R)$, we have, without approximation,

$$c_s = c_0[1+r/(F_0-r)]. \qquad \ldots\ldots(12)$$

The effect of the short circuit on the value of C_0 for the primary circuit will be seen to be insignificant.

Under the ideal conditions of negligible self-inductance of the rotor and negligible times of break and short circuit at commutation, as the rotor revolves with a frequency n and period T, the condition for balance in the detector is readily seen. For in a time $T/2$ the value of M changes from a negative maximum, $-M_0$, just after commutation to a positive maximum, M_0, just before the next commutation. The flux change, ϕ, in the detector circuit due to the sweep in the time $T/2$ is thus $2CM_0$, for there is no flux change due to the commutation or to the self-inductance, λ. The flux change, ψ, in the primary circuit in the time $T/2$, is algebraically zero, because the change $2c_0M_0$ between the commutations is neutralized by the same amount of opposite sign on actual commutation, whilst the self-inductance, μ, is without influence in half a period. Thus the integral of (5) gives

$$Q = \int_0^{T/2} c \, . \, dt = \frac{ER}{Y} \, . \, \frac{T}{2} - \frac{X+R}{Y} \, . \, 2C_0M_0, \qquad \ldots\ldots(13)$$

and for zero aggregate quantity of electricity through the detector over any number of half cycles, we have

$$R = 4 \, . \, nM_0 \, . \, C_0(X+R)/E, \qquad \ldots\ldots(14)$$

which, by virtue of (6), (7) and (8), is more usefully expressed

$$R=4\,.\,nM_0[1+c_0R/E] \qquad \ldots\ldots(15)$$

$$R=4\,.\,nM_0[1+R^2/(FX+FR+XR)]. \qquad \ldots\ldots(16)$$

The balance when the rotor is inductive and is short-circuited at commutation

A complete period of rotation is made up of the time $t=t_1+t_2$ of the two approximately equal short circuits which cover the commutations, and of the two longer intervals of combined duration $(T-t)$.

During the time $(T-t)$, the motional flux changes are $\phi=4C_0M_0$ in the detector circuit and $\psi=4c_0M_0$ in the primary circuit, producing flows in the negative direction in the detector; there is no resultant flux change due to any of the self-inductances, μ, λ and L. Thus, from (5) and (8), the quantity of electricity which passes is

$$Q_{T-t}=c_0(T-t)-4\,.\,C_0M_0\,.\,(X+R)/Y-4\,.\,c_0M_0R/Y. \qquad \ldots\ldots(17)$$

During the time t, the current through the detector is c_s of (11) and (12), and during the initial portions of each short circuit, the flux c_0L through the isolated rotor decays through r, being finally destroyed by the momentary break at commutation. The destruction of this flux is without influence on the detector. On leaving each short circuit, the current c_0 is re-established in the rotor in the contrary direction. Thus, as a result of each short circuit and commutation, a flux change, $\phi=c_0L$, having negative influence, is produced in the detector circuit, while a flux change, $\psi=2c_0M_0$, having positive effect on the detector, is produced in the primary circuit. A small portion of this last effect on the primary may reach the detector while the secondary circuit is reduced by r, i.e., with Y_s of (10) replacing Y of (6); but the effect of this can be shown to be negligible, not affecting the expression for R by more than a few parts in 10^9. The self-inductance, λ, is without effect on the quantity of electricity conveyed by the current c_s. Thus from (5) and (12) the quantity of electricity which passes the detector due to both short circuits and commutations is

$$Q_t=c_0\left[1+\frac{r}{F_0-r}\right]t-\frac{X+R}{Y}\,.\,2c_0L+\frac{R}{Y}\,.\,4c_0M_0. \qquad \ldots\ldots(18)$$

Adding (17) and (18), we have for the quantity of electricity traversing the detector in a whole period

$$Q_T=c_0T-\frac{X+R}{Y}\,.\,4C_0M_0+\frac{c_0rt}{F_0-r}-\frac{X+R}{Y}\,.\,2c_0L. \qquad \ldots\ldots(19)$$

The condition that the last two terms shall have no resultant effect is

$$t=2L(1/r-1/F_0), \qquad \ldots\ldots(20)$$

and, if this condition is satisfied, Q_T will be zero when the first two terms are equal, that is, when R has the value given by (14), (15) and (16). Under actual experimental conditions, the correcting term of (16) is of the order of 3 parts in 10^6, so that the basic expression, (1), is sufficiently accurate.

Attainment of balance between the effects of short circuit and self-inductance

The last two terms of (19) have, in actual experiment, an order of magnitude of some 1·2 parts in 10^4 when compared with the equal fundamental terms in T and M_0. By equating these minor terms to an accuracy of 2 % the combined error due to short circuit and self-inductance will be less than 3 parts in 10^6.

This is done by carrying out a compensation test with the arrangement seen in essence in figure 7 below, in which the rotor is switched into a simple circuit of some 1800 ohms, containing a 2-volt accumulator. The e.m.f. of about a volt drawn off half this resistance is balanced on a suitable potentiometer with the rotor at rest and unshunted. The rotor is then set spinning at the standard speed of rotation and the potentiometer balance examined. By adjustment of r_2, the dynamical and statical potentiometer balances are made identical. The potentiometer, under the actual conditions of test, was a measurer of the average current under the dynamical conditions to an accuracy of 1¼ %, and the effect of the earth's field, which is neutralized by field coils, has less than 5 % of its sensitivity in the neutralizing test.

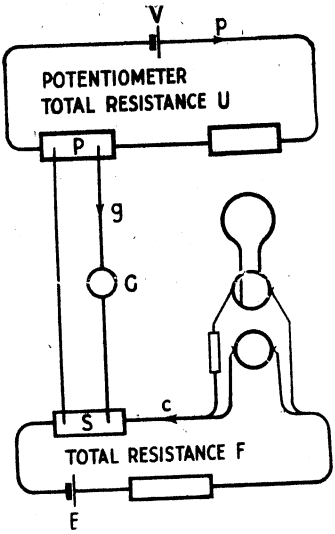


Figure 7. Essence of the compensation test.

If, in this compensation test, F_s is the total resistance of the rotor circuit, $r_s = r_1 + r_2$, the resistance cut out by short circuit, and $\bar{c}$, the statical current balanced off S on the potentiometer, we have, on attaining identity of statical and dynamical balances,

$$\bar{c} \,.\, T = \int_0^T \bar{c}\,dt = \bar{c}T + \bar{c} r_s t/(F_s - r_s) - \bar{c} \,.\, 2L/F_s, \qquad \ldots\ldots(21)$$

whence $$t = 2L(1/r_s - 1/F_s). \qquad \ldots\ldots(22)$$

Now in the main experiment, the effective resistance F_0 of the detector circuit is greater than F_s; the balancing value r_s in the compensation test should accordingly be raised to r, where

$$1/r = 1/r_s - 1/F_s + 1/F_0. \qquad \ldots\ldots(23)$$

Thus, with $r_s = 25$ ohms, $F_s = 1800$ ohms, $F = 8000$ ohms, then $r = 25{\cdot}27$ ohms.

The experimental value of r needed for compensation in the main experiment, corresponding to any value of the mean angle θ of the two short circuits, is in good agreement with that calculated, thus:—Let θ_1, θ_2 in degrees of arc be the two angles of short circuit, covered in the times t_1, t_2 during the period T. Then

$$t_1 + t_2 = t = T(\theta_1 + \theta_2)/360 = T\theta/180, \qquad \ldots\ldots(24)$$

whence, from (20),

$$2L(1/r - 1/F_0) = T\theta/180, \qquad \ldots\ldots(25)$$

and, since F_0 is large, $L = 0{\cdot}0385$ H., $1/T = 12{\cdot}5$, we have approximately

$$r\theta = 173. \qquad \ldots\ldots(26)$$

Experiments have been performed with various values of θ ranging from $5^\circ{\cdot}2$ to $8^\circ{\cdot}65$ and corresponding values of r of from $33{\cdot}5$ ohms to $20{\cdot}2$ ohms.

Use of the potentiometer to measure average current in the special circumstances

Applying Kirchhoff's laws to the three loops of figure 7, and using the notation of the diagram, we have, in general, for the current through the galvanometer when the rotor is unshunted,

$$g = [VP/U - (E - Lc)S/F]/Z, \qquad \ldots\ldots(27)$$

where $$Z = G + P - P^2/U + S - S^2/F, \qquad \ldots\ldots(28)$$

and, on securing static potentiometer balance,

$$PFV = SUE \qquad \ldots\ldots(29)$$

On spinning at standard frequency, F is reduced to $(F - r)$ and Z is reduced to a new value Z' during the times of short circuit of total duration t; a current of value

$$g' = -E \cdot \frac{S}{F}\frac{r}{F-r} \cdot /Z' \qquad \ldots\ldots(30)$$

then flows through the galvanometer.

During each of the intervals between the short circuits, the current through the galvanometer is for the most part zero, but at the beginning of these intervals there is a deficit in c due to the establishment of current in the inductive rotor and a flow of electricity through the galvanometer of quantity:

$$Q_0 = ESL/F^2Z. \qquad \ldots\ldots(31)$$

Thus, for balance still to hold under the dynamical conditions at the statical setting of the potentiometer,

$$tg' = 2Q_g, \qquad \ldots\ldots(32)$$

i.e.

$$\frac{t}{2L} \cdot \frac{Fr}{F-r} = \frac{Z'}{Z}, \qquad \ldots\ldots(33)$$

or

$$t = 2L(1/r - 1/F)(1-\alpha), \qquad \ldots\ldots(34)$$

where

$$\alpha = r \,.\, S^2/F(F-r)Z. \qquad \ldots\ldots(35)$$

Hence, if $S = F/2$, we have, approximately

$$\alpha = r/4(P - P^2/U + G + F/4), \qquad \ldots\ldots(36)$$

and with $r = 25$ ohms, $G = 25$ ohms, $P = 50$ ohms, $U = 100$ ohms, $F = 1800$ ohms and $S = 900$ ohms, we have $\alpha = 1/80$. A galvanometer of some 500 ohms would reduce α to less than 1/150 and be otherwise advantageous.

The earth's vertical field is neutralized in this test as in the main experiment, and the dynamical balance is taken for both directions of a reversing key at the rotor brush terminals. It is advantageous, however, that any small average e.m.f. dE, due to imperfect neutralization, should have smaller effect than in the test for neutralization. The current through G due to dE in the potentiometer test is

$$g = dE \,.\, S/FZ. \qquad \ldots\ldots(37)$$

With $S = F/2 = 900$ ohms, this is only some 5 % of the effect when $S = F$, and can be shown to be about 4½ % of the effect in the direct test for neutralization.

§ 4. THE EXPERIMENTAL ARRANGEMENT

The general form and lay-out of the main apparatus will be gathered from figures 8, 9 and 10. The twin field coils A, B, (figure 8) were constructed of dexonite. The channels of radial depth and axial breadth, 3·6 cm., were each filled with 504 turns of d.s.c. copper wire of S.W.G. 16, the mean diameter of the windings being 32·8 cm. The coils were firmly clamped, after symmetry tests, with the planes of the windings horizontal at a mean distance of separation of 8·1 cm., sufficient to permit the passage of the rotor shaft, and their series connection could readily be switched from conjunction to opposition.

In order that the maximum mutual inductance between the rotor and the field coils should be capable of continuous variation by ± 1 %, five additional turns were wound round the outer central portions of the coils; each twin pair of turns, one on each field coil, contributed some 37 μH., the leads being brought, as shown at J in figure 10, to separate dial terminals, 0–5; furthermore, two turns (Cn) of constantan wire (S.W.G. 18) were also wound round each field coil and joined in series with a 1-ohm dial resistance-box in tenths so as to be in parallel with suitable rheostats, Rh_2, Rh_3, and thus provide, with fine adjustment, a continuous range of from 0 to 40 μH. This system, which is essentially that described by Astbury (1938) in the construction of a standard mutual inductance, was most satisfactory, the impurity effect being negligible.

The complete rotor of the type described in § 2 is seen in figure 8. Each of the coils C and D was wound on a former having a channel of width 1·1 cm. and of internal diameter 11·1 cm., with 308 and 310 turns respectively of d.s.c. copper wire of s.w.g. 26; they were connected by twin bell flex which passed through the oak bed and hollow brass spindle to the insulated terminals a, b, c, d (figure 9) on the commutator M. After adjustment and fixing of the field coils for symmetry with respect to the horizontal axis of rotation, each rotor coil was separately adjusted in normal displacement from this axis, with the aid of flat oak distance pieces, so as to obtain a symmetrical and maximum flat of mutual inductance with the plane of its windings horizontal and parallel to those of the field coils; the rotor coils were then connected at the commutator in series and conjunction and had then a maximum inductance with the field coils of 20,032·7 μH. (nominal), which could be varied with the aid of additional windings to any value between 19,806 μH. and 20,258 μH. Although the rotor was of larger size than that described in § 2, the "flat" of mutual inductance over a range well exceeding the largest possible short circuit was so highly satisfactory that it was decided, in view of the imminence of war, to proceed at once with absolute measurement; our intention of later determining the precise wave form of the

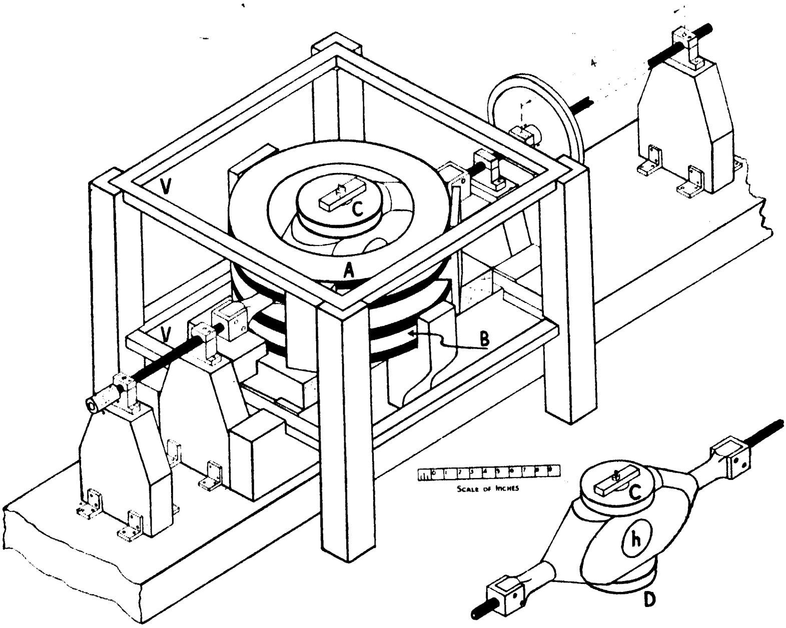


Figure 8. The field coils and rotor.

generator was in the end frustrated. The resistance of the rotor across the break terminals was 17·5 ohms and its self-inductance 38,510 μH.

The commutator M and the shorting rotor N, with the brushes H, are seen in figure 9. The ebonite-filled gaps between the brass halves of the commutator are about 1° of arc. The final form of shorting rotor consisted of an ebonite cylinder 6 cm. long and of 5 cm. diameter, through the centre of which passed a stout brass tube fitted with set screws which served to secure the whole to the hollow shaft conveying the rotor leads to the commutator. The ebonite cylinder was covered with a tightly fitting brass tube of 5·6 cm. external diameter which, between the binding screws near the edges, had two straight cuts 4 mm. wide, the one parallel to the axis of the cylinder and the other slightly inclined thereto, so as to divide the whole into two portions, the larger of which varied in angle over the length of the cylinder from 183° to 190°, and the smaller from 159° to 152°. After filling the gaps with ebonite and rounding true, the rotor yielded the desired range of short circuit of some 4° to 9° between diametrically opposite brushes, the variation being obtained by merely sliding the rotor along the shaft and re-clamping. The spring brass brushes, with their light steel springs and sponge-rubber dampers, had slight adjustment in all directions;

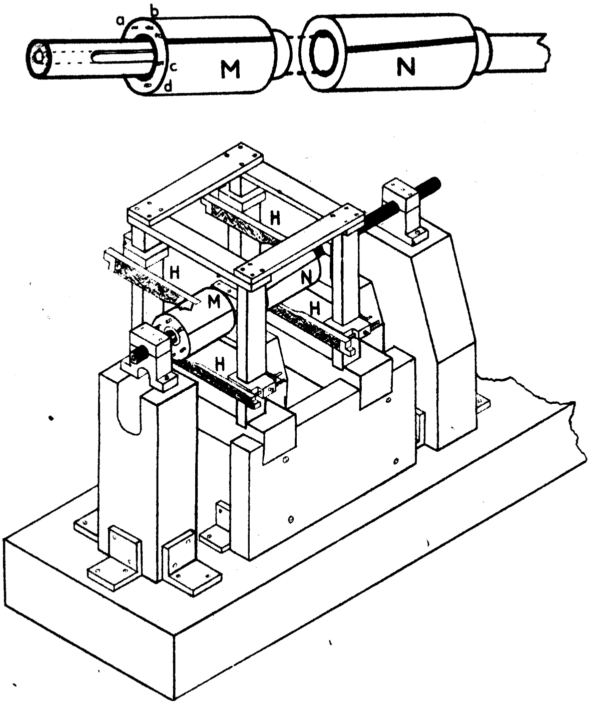


Figure 9. The commutator and shorting rotor.

they were the outcome of many trials and were most satisfactory, being preferable for constancy in angle of short circuit, even contact, and clear break at commutation, to the fiddle-bow brushes used in the preliminary investigation (Balls, 1938).

The horizontal axis of rotation lay in the magnetic meridian, and the earth's vertical flux through the rotor was neutralised by the large horizontal coils V, V, of sides 49 cm. by 55 cm., placed in the equivalent Helmholtz position and conveying a current of about 0·22 amperes.

The drive and speed regulation was similar to that already used by us and described elsewhere (Nettleton and Balls, 1935; Balls, 1938.) In this case, however, a special valve-maintained tuning fork with very fine adjustment of frequency around 375 was used with valve amplification to supply current to the coils of a 30-toothed-wheel synchronous motor as well as to a unilateral neon lamp which illuminated a 30-line stroboscopic disc on the shaft. At first, the fork frequency was determined by taking the revolutions with two twin telephone counters over prolonged periods of synchronization of the motor, detached from the main shaft in a manner already described (Balls, 1938.) Later, the fork was accurately adjusted to a frequency of 375 at 18°·7 c. by use of the 1000-cycle note broadcast by the N.P.L., which was amplified to light a second neon lamp and so obtain, through a low-power telescope, a stroboscopic view of an 80-line disc on the shaft. Unfortunately, this note was only broadcast monthly and, owing to interference with other lines, it was not possible to have the 1000-cycle note of the P.O. Research Station transmitted to us over the telephone. For intermediate checks we had to make use of the daily long-wave transmission of Deutschlandsender which, owing to local interference, gave a pattern less clear and more difficult to follow than that given by the N.P.L. signal. During the course of two months, during which all the more important tests were made, the frequency of the fork remained so constant that the only change necessary to maintain the standard speed of revolution was a change in temperature of the fork from 18°·7 c. to 18°·9 c.

The chief features of the electrical circuit are shown in figure 10. The standard mutual inductance was 10,007·5 ± 0·1 c.g.s. μH. at 10 cycles and was equivalent under D.C. to a reading of 10,011·2 nominal μH. on the Campbell inductometer. Thus the {rotor, field-coil} mutual inductance when balanced against the standard together with the inductometer at a stud reading of 10,000 and m divisions on the scale is $20{,}003{\cdot}8+m$ true μH. The circuit permits of tests for interaction between the inductometer and the standard, which was carefully eliminated.

A useful expedient is the key P_{11} in one of the brush connections of the shorting rotor, with leads, V_1, V_2, therefrom to wander plugs and the lead V_0 from r_2 serving for connection with an ohmmeter. This permits of individual tests of the contacts of the commutator and shorting rotor and enables the latter to be set symmetrically to cover the breaks in the former when the planes of all coils are parallel. Moreover, in measuring the sweeps of inductance between the rotor and field coils, the leads, V_0, V_1, enable the former to be set sharply in turn at the four boundaries of the short circuits, which are readily put out of action while the inductometer readings are taken.

The central auxiliary circuit, including R, S and W, serves, with slight changes, for the estimation of θ, the mean angle of short circuit, for the potentiometer test for the compensation of the self-inductance of the rotor and for preliminary observation of the main balance before switching over to the more delicate detectors. The angle θ, though not required directly, is of interest and enables the resistance r to be adjusted by virtue of equation (26) and by means of the box r_2 prior to an experimental compensation test; it is found by switching on the accumulator E_2 through the appropriate keys to the shorting rotor with r_2 infinite, and obtaining a full-scale static reading of 120 divisions on the micro-ammeter A_2 by adjustment of S and W. The rotor is turned to break contact

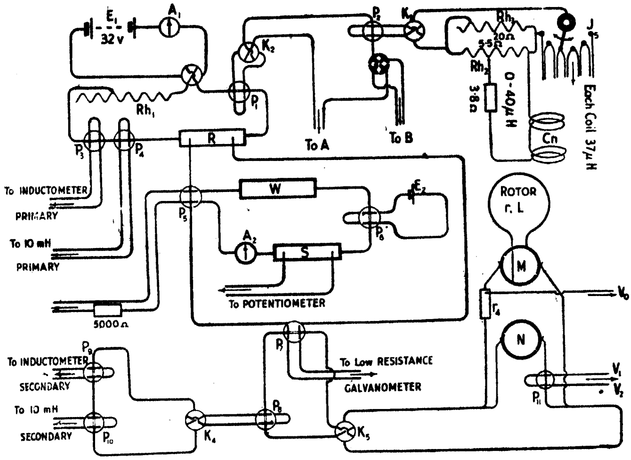


Figure 10. The experimental circuit.

through N, the voltage at E is increased m times without sensible change of resistance, and the deflection d is observed on the ammeter on spinning the rotor at any speed sufficient for steady reading. We have then $\theta = 3d/2m$. Potentiometer determinations of θ have also been made but are less simple.

In performing the compensation test, the effect of the earth's magnetic field is neutralized by adjustment of the current through the coils V, V (figure 8), as judged by the absence of deflection of a low-resistance galvanometer connected via the keys P_7, P_8 (figure 10) to the spinning rotor. The rotor is then brought to rest in any position in which the shorting rotor is inoperative, and a current is sent through it from the accumulator E_2 via W, R and S. With S = 900 ohms and W approximately the same, the e.m.f. across S of about a volt is balanced on the Weston-cell terminals of a thermoelectric potentiometer by adjustment of W and of the standard cell slide-wire. The motor is then set spinning at the standard speed and the potentiometer balance tested. If the dynamical balance reveals a loss of average current by fall in the e.m.f. across S, the effect of self-inductance prevails and r must be increased at r_2. When r is correct, the potentiometer reveals a clear gain at slow speeds which diminishes as the rotor gathers pace and which vanishes on synchronization. the dynamical balance being identical with the statical for either direction of the reversing key K_5. The value of r, easily sensitive to 2 %, was then raised by 0·2 ohms to 0·5 ohms in accordance with equation (23).

In the main experiments, the rheostats Rh_1 are adjusted for a current of from 0·6 to 0·8 amperes to flow from the accumulators E_1 of 32 volts through the resistor R and the field coils, with the inductometer and standard primaries out of circuit. On spinning the rotor at standard speed, with the earth's field neutralized, approximate balance is obtained in the micro-ammeter A_2 before rocking over the key P_5 through some 5000 ohms on to, first, a pilot galvanometer and finally a high sensitivity galvanometer of period 21 sec. and resistance 3000 ohms. The coils of all the detectors were set symmetrically so that no deflection was produced by alternating current; the broadening of the line of light was inappreciable in the final balance with the sensitive galvanometer. The criterion of balance, attained by varying the mutual inductance with the coil switch J, if necessary, and the rheostats Rh_2, Rh_3, is no permanent movement of the spot of light on reversing the key K_1, a procedure which keeps constant any residual effect due to the earth's field and other causes. No detectable change in the balance was ever observed if taken in this way with both the keys K_2 and K_5 reversed. The balance is confirmed by repeated reversals on the reverser K_1 at intervals of a minute or more, and any irregularity is usually overcome at once by applying a light brush moistened with petrol to the commutator or shorting rotor. Using a current of 0·8 amperes, and with the resistance of the detector circuit at 8000 ohms, the sensitivity was such that a full-scale deflexion of 75 cm. was obtained in one minute on the large scale common to both galvanometers at each reversal of the key K_1 when the mutual inductance was changed from the exact balancing value by 5·0 μH. The stability of the balance is closely bound up with the improvements made since the preliminary investigation in the accuracy of spinning between close bearings of the commutator and the shorting rotor, in the design of the latter and of the brushes, and in the use of non-abrasive cleaners and polishers.

Immediately after the balance, the rotor is brought to rest and, with the aid of key P_{11} and an ohmmeter set sharply on the four boundaries of short circuit, the inductometer readings are taken to 0·1 μH., using the same current as in the main experiment. For this test, the inductometer and standard primaries are switched into the main circuits at P_3 and P_4; P_5 and P_{11} are open and the rotor is joined through the two secondaries to a sensitive low-resistance galvanometer of 3 sec. period. By taking the mean of these four readings in estimating M_0, not only have we a true measure of the actual flux operative, but error due to a fixed mutual inductance between the field coils and the portion of the secondary circuit between the commutator brushes and the reversing key K_5 is eliminated.

§ 5. EXPERIMENTAL RESULTS

The data of a typical experiment are given in table 1 below:—

Table 1. Data of experiment 15

Speed of revolution at 18°·9 c.	$12{\cdot}5000_0$ rev./sec.
Angle of short circuit θ	$7^\circ{\cdot}4 \pm 0^\circ{\cdot}1$
Compensation resistance for shorting rotor r	$24{\cdot}0 \pm 0{\cdot}3$ ohms
Primary current C_0	0·8 amperes
Detector circuit resistance F	8000 ohms
Temperature of fork	19°·0 c.
Temperature of ohm	19°·6 c.
Mean scale inductometer reading m	$7{\cdot}3_0$ nom. μH.
Measured mutual inductance	$20011{\cdot}1_0$ true μH.
Correction for fork temperature	$-0{\cdot}2_2$ μH.
M_0	$20010{\cdot}8_8$ μH.
$R = 4M_0 \,.\, n \times 10^9$	$1{\cdot}00054_4$ ohm
R in international ohms at 19°·6 c.	$1{\cdot}00002_8$ ohm
Value of one international ohm	$1{\cdot}00051_6$ ohm

Table 2. Results of consecutive experiments

Exp.	θ° of arc.	r ohms	m nom. μH.	dT° c.	Temp. of ohm in °c	Value of ohm in c.g.s. units × 10^9	Value of ohm in internat. units	Difference
1	5·2	33·5	$6{\cdot}7_0$	0·0	18·8	$1{\cdot}00052_5$	$1{\cdot}00000_4$	$\cdot00052_1$
2	5·2	33·5	$6{\cdot}6_0$	0·0	19·0	$1{\cdot}00052_0$	$1{\cdot}00001_0$	$\cdot00051_0$
3	5·2	33·5	$7{\cdot}0_5$	0·2	19·1	$1{\cdot}00052_1$	$1{\cdot}00001_3$	$\cdot00050_8$
4	5·2	33·5	$6{\cdot}6_2$	0·1	18·9	$1{\cdot}00051_0$	$1{\cdot}00000_7$	$\cdot00050_3$
5	6·5	27·3	$7{\cdot}8_0$	0·5	19·4	$1{\cdot}00052_5$	$1{\cdot}00002_2$	$\cdot00050_3$
6	6·5	27·3	$7{\cdot}9_8$	0·4	19·3	$1{\cdot}00054_5$	$1{\cdot}00001_9$	$\cdot00052_6$
7	6·5	27·3	$7{\cdot}8_5$	0·4	19·3	$1{\cdot}00053_9$	$1{\cdot}00001_9$	$\cdot00052_0$
8	8·65	20·2	$7{\cdot}9_3$	0·5	20·0	$1{\cdot}00053_1$	$1{\cdot}00004_0$	$\cdot00049_1$
9	8·65	20·2	$4{\cdot}4_0$	−0·75	18·1	$1{\cdot}00049_3$	$0{\cdot}99998_3$	$\cdot00051_0$
10	8·4	20·7	$6{\cdot}2_5$	0·0	18·9	$1{\cdot}00050_3$	$1{\cdot}00000_7$	$\cdot00049_6$
11	8·4	20·7	$6{\cdot}4_5$	0·0	18·7	$1{\cdot}00051_3$	$1{\cdot}00000_1$	$\cdot00051_2$
12	7·4	24·0	$6{\cdot}6_2$	0·0	19·4	$1{\cdot}00052_1$	$1{\cdot}00002_2$	$\cdot00049_9$
13	7·4	24·0	$6{\cdot}4_2$	0·0	18·9	$1{\cdot}00051_1$	$1{\cdot}00000_7$	$\cdot00050_4$
14	7·4	24·0	$6{\cdot}5_5$	0·1	19·1	$1{\cdot}00050_7$	$1{\cdot}00001_3$	$\cdot00049_4$
15	7·4	24·0	$7{\cdot}3_0$	0·1	19·6	$1{\cdot}00054_4$	$1{\cdot}00002_8$	$\cdot00051_6$
16	7·4	24·0	$7{\cdot}7_8$	0·3	19·8	$1{\cdot}00054_6$	$1{\cdot}00003_4$	$\cdot00051_2$
						Mean difference . .		$\cdot00050_8$
One international ohm = 1·00051 ohm.								
17	7·4	24·0	$2{\cdot}3_0$	0·0	19·0	$1{\cdot}00030_5$	$0{\cdot}99981_0$	$\cdot00049_5$
18	7·4	24·0	$-3{\cdot}4_8$	0·0	19·0	$1{\cdot}00001_6$	$0{\cdot}99951_0$	$\cdot00050_6$
19	5·2	33·5	$69{\cdot}1_0$	0·2	20·0	$1{\cdot}00362_8$		

In table 2 are first given the results of sixteen consective direct experiments spread over five weeks. Experiments 17 and 18 were performed with 5000 ohms and 2000 ohms, respectively, in parallel with the international ohm. Experiment 19 was performed between experiments 4 and 5 with an old laboratory german-silver Board of Trade ohm. m is the mean reading on the scale of the inductometer of the four boundary settings; dT is the temperature of the fork above that at which the speed of revolution was last confirmed at 12·5 rev./sec.

The consistency of the method, allowing for temperature changes, is shown by the extreme variation of 3·5 parts in 10^5 in 16 experiments. The tenth-ohm box linking the constantan coils Cn (figure 10) to the rheostats Rh_2, Rh_3 was often varied to prevent any fore-knowledge of the rheostat settings at which balance might be expected.

§ 6. CONCLUDING REMARKS

This paper is the last of the series of investigations on absolute measurement of electrical resistance carried out in recent years at Birkbeck College. This method in its present form far exceeds in accuracy any of our previous methods, in particular that of the rotating coil recently improved by one of us (Balls, 1938). This is because of the great sensitivity, not yet pushed to its limit, the adoption of the method to measure directly a resistance of an international ohm, and the need for accurate measurement only of a frequency of 1000 cycles and of a mutual inductance of very closely 20,000 μH. on a remarkable " flat." The 2% experimental compensation test is an important feature of the method, and the agreement here with theory is satisfactory in showing the insignificance of additional effects such as would arise from imperfect contact at the brushes. There would be little difficulty in attaining here an accuracy approaching 0·5 %, the chief advantage of which would be manifest in the stability of the main balance.

We regret that we are unable to reconstruct the apparatus and carry out further improvements—for we believe the method is capable of high precision. As it is, we had the satisfaction of knowing that the most accurate method of calibrating an ohm in our laboratories was by this rapid method of measuring its resistance directly in c.g.s. units.

§ 7. ACKNOWLEDGEMENTS

We express our gratitude to Prof. J. D. Bernal, M.A., F.R.S., for the help and encouragement that he has given us throughout this investigation. We also wish to thank Mr. W. Wilson, B.Sc., of the National Physical Laboratory, Dr. W. Ehrenberg, of Electrical and Musical Industries, Ltd., Hayes, and Mr. S. Baker, Mr. H. G. Bell and Mr. R. A. Dobb, of Birkbeck College, for the help and advice that they have given us.

REFERENCES

ASTBURY, N. F., 1938. *Phil. Mag.* **25**, 290.
BALLS, E. G., 1938. "The absolute measurement of electrical resistance by new rotating coil methods." Thesis (London).
CURTIS, H. L., 1937. *Elect. Measurements* (London : McGraw-Hill), p. 177.
LLEWELLYN, F. H., 1934. *Proc. Phys. Soc.* **46**, 839.
NETTLETON, H. R. and BALLS, E. G., 1935. *Proc. Phys. Soc.* **47**, 58.
ROSA, E. B., 1909. *Bull. Bur. Standards*, **5**, 499. See also GLAZEBROOK, *Dict. of Applied Physics*, vol. 2, p. 226.

NOTE ON ELECTRIC AND MAGNETIC DIMENSIONS

BY G. D. YARNOLD, M.A., D.PHIL.,
University College, Nottingham

Received 25 *August* 1941

ABSTRACT. Attention is drawn to the fact that the dielectric constant K and the magnetic permeability μ of a medium are essentially dimensionless. From the definitions given for these quantities the usual forms of the equations expressing the inverse-square laws of force are readily derived. The dimensions of the constant factor unity in each of these equations must not, however, be associated with K and μ.

THREE extremely interesting papers which have recently appeared in these *Proceedings* have raised again the vexed question of the physical dimensions of electric and magnetic quantities, and in particular of the dielectric constant K and the magnetic permeability μ of a medium. Thus, Brown (1940, 1941) maintains that both K and μ, being defined as ratios of quantities of similar dimensions, are themselves necessarily dimensionless; while Duncanson (1941) regards K and μ as possessing dimensions. The whole difficulty appears to arise from the confusion of two quite distinct quantities, viz., the constant which must always be included in the equation expressing an inverse-square law of force and the quantity known as the dielectric constant or the permeability of the medium, as the case may be. Of these, the former is found to possess dimensions, while the latter is by definition dimensionless. Benham, in the discussion on Brown's earlier paper (1940), noted this same point, but his remarks do not appear to have received the attention they deserve.

Adopting the usual nomenclature, the two inverse-square laws are expressed as follows:

$$\text{force between point charges } \textit{in vacuo} \propto \frac{q_1 q_2}{r^2},$$

$$\text{force between point poles } \textit{in vacuo} \propto \frac{m_1 m_2}{r^2}.$$

It should be noted that these laws have been established by experiment only in the case of charges or poles in air or vacuum. No person has any direct experimental knowledge of the forces between charges or poles " embedded " in any material medium other than a gas, and the extension of the laws to the general case rests entirely on other considerations.

The laws of force may be written as equations by introducing two constants a and b. Thus, referring still to vacuum,

$$F=a\frac{q_1q_2}{r_2} \quad \text{and} \quad F=b\frac{m_1m_2}{r^2}.$$

All that can be said at present about the dimensions of a, b, q and m is that aq^2 and bm^2 each have the dimensions of a force multiplied by the square of a distance, i.e. $[\mathrm{ML^3T^{-2}}]$ in terms of the usual three indefinables. It may be shown, however, that the product ab has the dimensions of the square of a velocity, i.e. $[\mathrm{L^2T^{-2}}]$. Thus, while there is still no information concerning the dimensions of a and b separately, it is clearly unjustifiable to treat them as dimensionless.

The introduction of a medium other than vacuum is most conveniently made in the case of the electrical problem by considering the change in the capacity of a condenser brought about by completely filling the space between the plates with the medium. The definition of the dielectric constant K of the medium is then

$$K=\frac{\text{capacity of condenser with dielectric}}{\text{capacity of condenser without dielectric}}.$$

It follows that the difference of potential between the plates of a condenser carrying fixed charges is reduced by the medium in the ratio $1/K$, and that the law of force for point charges embedded in the medium is

$$F=\frac{a}{K}\frac{q_1q_2}{r^2},$$

in which K is, by definition, a dimensionless number. It is assumed implicitly that the introduction of a small charge produces no disturbance of the electrical conditions in the medium ; in other words, the medium is treated from a strictly classical point of view.

In the magnetic problem, the medium is introduced by considering the mechanical forces acting on a small hypothetical test pole when placed in turn in an evacuated long, thin, cylindrical cavity and in an evacuated short, flat, cylindrical cavity in the medium, the axis of the long cavity being parallel to the direction of the magnetizing field, and the axis of the short cavity parallel to the direction of magnetization of the medium. The definition of the magnetic permeability μ of an isotropic medium is then

$$\mu=\frac{\text{mechanical force on test pole in short flat cavity}}{\text{mechanical force on test pole in long thin cavity}}.$$

Thus μ also is defined as a dimensionless number, and the law of force for point poles embedded in the medium may be obtained as follows:—

Consider an isotropic medium of permeability μ, containing two short flat cavities bounded by four parallel infinite planes A, B, C, D (figure 1). Let the flat cavities be connected by a long thin cavity parallel to the direction of the magnetizing field and perpendicular to the set of planes, and suppose all the cavities to be evacuated. A small positive magnetic pole may be transported slowly from a point Z between planes C and D to a point X between planes A and B, either via the long thin cavity or by the direct route through the medium. The work done by the operator is the same in either case. Hence the mechanical force acting on the pole when placed at any point Y in the medium is the same as that acting on the pole when placed in the long thin cavity. Now let that part of the medium lying between planes B and C be removed (figure 2). Since the planes A and D are infinite while their separation is finite, the mechanical force on the pole when placed at Y is the same as that which previously acted on it when placed at X or Z. Thus, by the definition of the magnetic permeability, the force acting on the small pole when embedded in the medium is equal to $1/\mu$ of the force acting on it in vacuum, provided the strength of the magnetizing field is unchanged. The general form of the law of force between point poles is thus

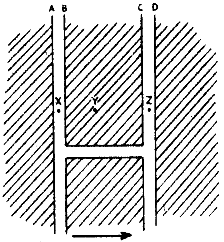


Figure 1.

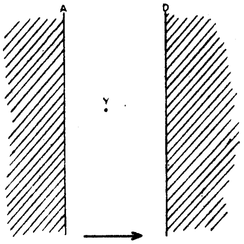


Figure 2.

$$F = \frac{b}{\mu}\frac{m_1 m_2}{r^2}.$$

It is unfortunate, in view of the formal similarity between the laws of force for electric charges and magnetic poles, that exactly analogous definitions cannot be adopted for K and μ. The reason for this no doubt lies in the fact that while on the atomic scale there is such a thing as an electric charge, there is no such thing as a magnetic pole. All modern work on magnetic phenomena shows that magnetic moment rather than magnetic pole is the fundamental entity, and that this in turn can be described in terms of moving electric charge.

The customary definitions of the electrostatic unit of charge and the unit of magnetic pole strength are designed to make the constants a and b each equal to unity when the forces are measured in dynes and the separation of the charges or poles in centimetres. The laws of force between point charges and point poles placed in a medium then reduce respectively to their usual forms:

$$F=\frac{q_1q_2}{Kr^2} \quad \text{and} \quad F=\frac{m_1m_2}{\mu r^2},$$

but it must always be borne in mind that these equations, as they stand, are dimensionally incorrect, since a constant equal to unity but possessing unknown dimensions is omitted from each. No corresponding ambiguity exists in the statement of the gravitational law of force, since in this case the constant G is retained and is recognized as possessing dimensions.

The confusion in electric and magnetic units arises because the dimensions proper to a and b respectively in the above equations are commonly associated with K and μ. In particular, it is thoroughly misleading to state that the square of the velocity of light is equal to the reciprocal of the product of the dielectric constant and the magnetic permeability of free space. In fact, both the dielectric constant and the permeability of free space are, by definition, dimensionless and equal to unity, and it is the product ab which may be identified with the square of the velocity of light, provided, of course, that the same system of electrical units is used throughout. It is felt that the general adoption of a notation similar to that employed above would reduce the present confusion.

From what has been said, it is clear that electric and magnetic dimensions cannot be expressed in terms of the three fundamental quantities mass, length and time only, and there appears to be much to commend the acceptance of electric charge as an additional indefinable. The dimensions of the constant a are, then, $[ML^3T^{-2}Q^{-2}]$ and those of b are $[M^{-1}L^{-1}Q^2]$, while K and μ are dimensionless. The dimensions of other quantities remain as given by Duncanson, with the exception that electric and magnetic induction have respectively the same dimensions as the corresponding field strengths.

The fact that precise dimensions have been assigned to the constants a and b does not in any way imply that the phenomena of electric and magnetic repulsion or the properties of free space are completely understood. The adoption of the four indefinables M, L, T and Q, each of which appears incapable of *adequate* representation in terms of the other three, is merely a matter of convenience, and makes possible a consistent system of electric and magnetic dimensions.

REFERENCES

Brown, 1940. *Proc. Phys. Soc.* **52**, 577.
Brown, 1941. *Proc. Phys. Soc.* **53**, 418.
Duncanson, 1941. *Proc. Phys. Soc.* **53**, 432.

DISCUSSION

Mr. S. Gill.* Dimensions of common physical quantities are usually written in terms of the fundamental quantities length, time, and mass, together with either permeability or permittivity. These systems are awkward owing to the two alternative ways of representing electrical and magnetic quantities, and the fact that fractional indices often occur.

---

* It seems worthy of record that Mr. Gill is only 15 years of age. (Ed.)

The choice of fundamental quantities is arbitrary, and I find that by taking as fundamentals length, time, electric charge and magnetic pole-strength, not only are fractional indices avoided, but a striking similarity between the dimensions of corresponding electrical and magnetic quantities is found. If any pair of corresponding magnetic and electrical quantities is taken, e. g. intensities of fields, the indices of length and time are found to be identical, while those of charge and pole-strength are interchanged. Moreover, in the case of " pure " physical quantities, the indices of charge and pole-strength are the same, so that interchanging has no effect.

	Length	Time	Charge	Pole-strength
Mass	−2	1	1	1
Momentum	−1	0	1	1
Force	−1	−1	1	1
Energy	0	−1	1	1
Moment of inertia	0	1	1	1
Angular momentum	0	0	1	1
Permeability	−1	1	−1	1
Field-strength (magnetic)	−1	−1	1	0
Magnetic induction	−2	0	0	1
Rate of change of flux	0	−1	0	1
Magnetic potential	0	−1	1	0
Permittivity	−1	1	1	−1
Electric field-strength	−1	−1	0	1
Electrical induction	−2	0	1	0
Electrical current	0	−1	1	0
Electrical P.D.	0	−1	0	1

This shows clearly the fundamentally complementary nature of electrical and magnetic phenomena, although in practice this is masked by the fact that they are manifested in such widely differing ways.

DIAMAGNETIC SUSCEPTIBILITY OF METHANE

By C. A. COULSON, M.A., Ph.D.,
University College, Dundee

Received 3 *October* 1941

ABSTRACT. Theoretical calculations of the diamagnetic susceptibility of methane CH_4 have been made on the basis of the molecular-orbital and electron-pair approximate wave functions. The results are compared with the self-consistent-field value, and it is shown that in this latter an error of about 25 % is due to the preliminary averaging of the positive charges on the four protons. The new values agree well among themselves, but are still considerably larger than the experimental value. The reasons for this are briefly discussed.

Very few theoretical calculations of the diamagnetic susceptibility of molecules have been reported in the literature. This is directly due to our ignorance of reliable wave functions. Indeed, the diamagnetic susceptibility depends immediately upon the wave function, and the various types of approximation employed in calculating this latter may be tested by a comparison of the predicted susceptibilities.

Interest in this problem has recently been revived by the stimulating work of Buckingham, Massey and Tibbs (1941). These authors have dealt very fully with the methane molecule, using the self-consistent-field method (s.c.f. method) of Hartree. But in order to employ the customary technique of the s.c.f. approximation, it was necessary to average the four protons round the surface of a sphere with the carbon atom as centre; in this manner they procured the requisite spherical symmetry. There is probably no great error involved in this averaging process, especially with molecules so nearly spherical as CH_4, even though, as these authors explain, on account of the relatively slow motion of the nuclei compared with that of the electrons, it is a more drastic approximation than that used in the normal s.c.f. method for atoms. It is, however, desirable that we should test this averaging process as fully as we can, especially in view of its possible application to other molecular systems, and in this note, therefore, we report alternative calculations of the diamagnetic susceptibility, using molecular wave functions in which no averaging of the protons has taken place.

The formula for the diamagnetic susceptibility is (Van Vleck, 1932)

$$\chi_{\mathrm{mol}} = -\frac{Ne^2}{6mc^2}\Sigma\overline{r^2} + \tfrac{2}{3}N \sum_{n' \neq n} \frac{|\, m^\circ(n', n)\,|^2}{h\nu_{n'n}},$$

in which N is Avogadro's number and the other symbols have their usual meanings.

The summation in the first term is over all the electrons in the molecule, and in the second term over all non-diagonal matrix components of the electric moment. In the case of spherical symmetry this latter term is identically zero, and it is very unlikely to be large for methane, if we measure r from the carbon nucleus. We shall accordingly neglect this contribution and thus follow the course adopted by Buckingham, Massey and Tibbs. Our problem is therefore reduced to a calculation of the mean value of r^2 for each electron present.

Our first molecular wave functions are the molecular-orbital wave functions (m.o. approximation) determined by the present writer some years ago (Coulson, 1937). According to these calculations, the ten electrons are divided into three groups. The first group contains what are essentially the two K electrons of carbon, whose normalized wave functions are

$$\psi(C:1s)=\sqrt{(\gamma^3/\pi)}e^{-\gamma r},\qquad \gamma=5\cdot 7.$$

We are using atomic units throughout. The second group, called by Mulliken [s], has the complete symmetry of the molecule and contains two electrons with normalized wave functions:

$$\psi[s]=N_s\{8\cdot 22\,\psi(C:2s)+\psi(H_a)+\psi(H_b)+\psi(H_c)+\psi(H_d)\},$$

where

$$\psi(C:2s)=\sqrt{(c_s{}^5/3\pi)}re^{-c_s r},\qquad c_s=2\cdot 98,$$
$$\psi(H_a)=\sqrt{(1/\pi)}\ e^{-r_a},$$
$$N_s=0\cdot 10145.$$

The third group, called by Mulliken [t], holds two electrons in each of three orbitals, a typical member being

$$\psi[t_x]=N_t\{1\cdot 70\,\psi(C:2p_x)+\psi(H_a)+\psi(H_b)-\psi(H_c)-\psi(H_d)\},$$

where

$$\psi(C:2p_x)=\sqrt{(c_p{}^5/\pi)}xe^{-c_p r},\qquad c_p=1\cdot 62,$$
$$N_t=0\cdot 3259,$$

and the x axis bisects the line joining H_a and H_b. The other two members of this series are obtained by an obvious cyclic interchange.

It is not difficult to evaluate $\overline{r^2}$ for each group of wave functions. The integrals involved, some of which are fairly complex, may all be obtained from tables of such integrals prepared by the writer (Coulson, 1941). The mean values, in atomic units, are given below:

Wave function	$\psi(C:1s)$	$\psi[s]$	$\psi[t]$
$\overline{r^2}$	0·092	1·527	5·018

It will be noticed that the [s] type orbit is much more concentrated than the [t] type, and, in fact, almost all the susceptibility arises from the six [t] electrons. The value of χ deduced from these mean values is

$$\chi_{\text{mol}}=-26\cdot 6\times 10^{-6}\text{ e.m.u.}$$

Before we discuss this result we must deal with the second approximation. This is the electron-pair method (e.p. method) of Heitler and London. The electrons are taken in pairs and exchange is allowed for between the two members of a pair, but not between members of different pairs. There are two inner K-shell electrons similar to those described above in the m.o. approximation, and then four pairs (one pair for each C-H bond). Each of these pairs represents a homopolar bond and has a wave function such as

$$\Psi(\sigma)=N_\sigma\{\psi(\sigma:1)\psi(H_a:2)+\psi(\sigma:2)\psi(H_a:1)\},$$

where N_σ is a normalizing constant,

and $\psi(H_a)=\surd(\beta^3/\pi)\ e^{-\beta r_a},$

$$\psi(\sigma)=\tfrac{1}{2}\psi(C:2s)+\frac{\surd 3}{2}\psi(C:2p).$$

$\psi(\sigma)$ is, of course, the Pauling tetrahedral orbit, and

$$\psi(C:2s)=\surd(\alpha^5\beta^5/3\pi)re^{-\alpha\beta r},$$

$$\psi(C:2p)=\surd(\alpha^5\beta^5/\pi)xe^{-\alpha\beta r}.$$

The value of α is taken to be that given by Slater (1932), viz., $\alpha=1\cdot625$. We have, however, not retained the Slater exponent α unaltered, but have included, both in $\psi(\sigma)$ and $\psi(H)$, an extra factor β. This is introduced because, as Coulson and Duncanson (1941) have shown, some such factor is necessary to allow for different screening in passing from atomic to molecular wave functions. A value $\beta=1$ corresponds to the normal Slater functions for atoms, but a value $\beta=1\cdot1$ is probably more accurate for molecules. For the present, then, we retain $\beta=1\cdot1$, and later we shall discuss the effect of varying β between $1\cdot0$ and $1\cdot2$. We do not require to introduce this factor for the $1s$ electrons, which are substantially unaffected by molecular formation.

Using these wave functions, we find that, in atomic units,

$$\overline{r^2(1s)}=0\cdot092,\qquad \overline{r^2(\sigma)}=4\cdot313.$$

The resulting value for χ_{mol} is $\chi_{\text{mol}}=-27\cdot7\times10^{-6}$ e.m.u. The various values of χ_{mol} are collected in the table.

Table. Values of the diamagnetic susceptibility

Molecular-orbital	$-26\cdot6\times10^{-6}$ e.m.u.
Electron-pair ($\beta=1\cdot1$)	$-27\cdot7\times10^{-6}$ e.m.u.
Self-consistent-field	$-33\cdot2\times10^{-6}$ e.m.u.
Experimental value (Bitter, 1929)	$-12\cdot2\times10^{-6}$ e.m.u.

We may make several deductions from this table. First, the agreement between the m.o. and e.p. calculations is remarkably good. It is true, unfortunately, that the m.o. method allows for ionicity but not for exchange, whereas the e.p. method allows for exchange but not for ionicity, and in each case the neglected factor would reduce the numerical value.

The s.c.f. method of Buckingham, Massey and Tibbs may be compared with the m.o. method, which it resembles very closely, since they both allow for ionicity but not for exchange. It is seen that the s.c.f. value is about 25 % larger than the m.o. value. This may be interpreted as showing that the effect of averaging the four protons is to give too diffuse a charge cloud with too large a value of $\overline{r^2}$. We may expect a similar kind of result for other molecules of this type, and the averaging process will probably need some modification if it is to be applied to heavier molecules, such as CCl_4.

The values obtained in this paper with the m.o. and e.p. approximations, though less than that obtained from the s.c.f. approximation, are still about twice the experimental value; this factor, which is somewhat larger than that usually found with atoms, must be due chiefly to approximations in the wave functions. Indeed, there is every reason to believe that, as with atoms, if we could develop more accurate wave functions we could reduce the error very considerably. It is a pity that satisfactory molecular wave functions are so laborious to compute. In the present type of problem the presence of r^2 in all the integrands reveals that the largest influence comes from the outer regions of the molecule in which the detailed form of the wave functions is known with least reliability.

In conclusion it may be wondered whether the value of χ_{mol} found by the e.p. approximation is sensitive to the presence of the scale factor β, and, if so, whether there is much significance to be attached to the corresponding value in the table. We have therefore made equivalent calculations, taking $\beta = 1\cdot0$, $1\cdot1$ and $1\cdot2$, and the results are shown below:

$\beta =$	$1\cdot0$	$1\cdot1$	$1\cdot2$
$\chi_{\text{mol}} =$	$-31\cdot0$	$-27\cdot7$	$-25\cdot3 \times 10^{-6}$ e.m.u.

The appropriate value of β is almost certain to lie within this range. Arguing by analogy from the case of molecular hydrogen H_2, we should expect β to lie between $1\cdot1$ and $1\cdot2$. This table shows that any plausible alteration in the value of β is inadequate to explain the difference between theory and experiment. This must be due, not so much to inaccuracy in the finer details of the wave functions, as to the coarser approximations, in which we neglect, on the one hand, ionic terms, or, on the other hand, the full determinantal symmetry of the molecular wave function.

REFERENCES

Bitter, 1929. *Phys. Rev.* **33**, 388.
Buckingham, Massey and Tibbs, 1941. *Proc. Roy. Soc.* A, **178**, 119.
Coulson, 1937. *Trans. Faraday Soc.* **33**, 388.
Coulson, 1941. *Proc. Camb. Phil. Soc.* (To appear shortly.)
Coulson and Duncanson, 1941. *Proc. Camb. Phil. Soc.* (To appear shortly.)
Van Vleck, 1932. *Electric and Magnetic Susceptibilities*, p. 275 (Oxford: Clarendon Press).

THERMOSTATS EMPLOYING EXTERNAL SURFACE CONTROL

By V. D. HOPPER, M.Sc., A.Inst.P.,
University of Melbourne

Received 28 *October* 1941; *Communicated by Prof T. H. Laby, F.R.S.*

ABSTRACT. An introduction is given by T. H. Laby, F.R.S., in which the principles involved in the production of constant temperature are discussed. These principles have been applied to the design of temperature-controlling systems using an external resistance thermometer to control the heat supply, which may be radiant heat or heated air. The hunting in the system in each case is less than 1 second. Control of the temperature of apparatus to better than 0·01° c. has been obtained by using these methods.

INTRODUCTION.

By T. H. Laby, F.R.S.

In scientific experiments and industrial operations it is often necessary to control the temperature of an instrument or of chemical or other plant. Principles which apply to such control are discussed below, and methods are described in which the instrument whose temperature is controlled is heated by radiation or warmed air, the heating supply being turned on and off by a resistance thermometer which is wound on the surface of the instrument. The time-lag of the thermostat is small, and the temperature fluctuations at the surface penetrate only to a small depth and the internal temperature is constant.

Mr. V. D. Hopper and the writer (Laby and Hopper, 1939), in measuring the electronic charge, devised a method to maintain at a constant temperature the air in which the fall of oil drops was under observation, and it proved most effective for that purpose. The resistance thermometer formed an arm of a direct-current bridge. The light spot of the galvanometer connected to the bridge operated a photoelectric cell, and the amplified current from that cell actuated a relay which turned on and off two lamps whose radiation heated the instrument.

Mr. Hopper at this stage was appointed to the Physics Laboratory of the University of Western Australia, where he has made tests of the radiation and warm-air thermostats. He has devised a bridge circuit and used an amplifier which employs A.C. current.

The principles which apply to the maintenance of an instrument at a constant temperature a few degrees above the maximum ambient temperature may be briefly analysed. It is noted that the ambient temperature not only changes daily and seasonally, but that, superimposed on these comparatively long-period temperature variations, there are fluctuations of much shorter period. The systems considered lose heat by conduction, convection and radiation to the surrounding bodies. Newton's law of cooling no doubt applies to the losses of heat, as the excess temperatures which determine them are small.

Heat is supplied to the system whose temperature is to be kept constant, at a rate controlled by the thermometer. In an ideal thermostatic system its rate of loss at every instant is equal to its rate of gain of heat, that is, there is no time-lag in the control.

In this introduction the following may be mentioned :

(1) Time-lag in the control of the heat supply.
(2) Depth of penetration of temperature fluctuations.
(3) Thermal shielding.
(4) External heating and thermometry.

Time-lag. In any thermostat there is a lag in time between a change in the temperature of the thermostat and the change in the heat supply. A general theoretical study of this time-lag in control systems has been given by Callendar, Hartree and Porter (1936). These authors show that for stable control of temperature (1) the supply of heat to the thermostat must return quickly to its normal value, and (2) there must be positive damping in the controlling system. This means that the thermometer which controls the heat supply must follow temperature changes rapidly and regulate the heat supply with a minimum delay.

A control with an inherently small time-lag is described later. It is a resistance thermometer of fine electrically insulated wire wound on the surface of the instrument to be controlled and connected to a Wheatstone bridge. The system, consisting of the thermometer, bridge and galvanometer, has a time-lag determined by the thermal capacity of the resistance thermometer and the period of the galvanometer, but these can readily be made small.

Depth of penetration of temperature fluctuations. Numerical values of the depth to which fluctuations of temperature penetrate into solids are given in the table on p. 59. It will be seen that fluctuations of temperature of short period penetrate thermal insulators only to a small depth. A temperature fluctuation of 1 sec. period has an inappreciable effect at a depth of 2 mm. below the surface of a sheet of cork. While the penetration is much larger for a good conductor, an enclosure constructed of alternate layers of good and bad conductors is an effective shield against temperature fluctuations.

Thermal shielding. In an enclosure of alternate layers of conductors and non-conductors, the non-conductor eliminates fluctuations of short period, and the layer of good conductor brings all points inside it to the same temperature.

Heating and thermometry at external surfaces. In some types of thermostat, the heater is external to the thermostat and in others an internal heater is used, the thermometer which regulates the supply of heat being placed inside the thermostat. With the thermometer in this position, temperature changes, at least sufficient to actuate the thermometer, occur in the region from which it is the purpose of the thermostat to exclude any temperature change.

If a mercury thermometer is used it cannot readily be placed outside the thermostat, but it may be recalled that Griffith constructed a thermostat enclosed by its mercury steel thermometer. If a resistance thermometer is used, it can be readily wound on the external surface of the instrument whose temperature is to be controlled. If this surface is a good conductor it is only necessary to cover a part of that surface.

The thermostat can be heated by radiation or a stream of heated air. The time-lag can be made small, and the surface fluctuations in temperature are damped out before penetrating far into the thermostat.

---

§ 1. CONTINUOUS AND DISCONTINOUS METHODS OF CONTROL

The temperature-control system employed consists essentially of a heater A (see figure 1), a means B of transmitting heat from the heater to the apparatus, a thermometer C which measures and controls the entry of heat into the shielding system D, and the object E whose temperature is controlled.

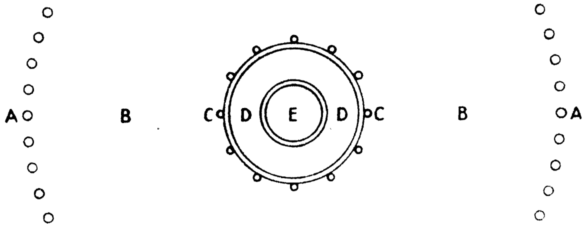


Figure 1. Temperature-controlling system.
A, heaters ; B, intervening medium ; C, thermometer ; D, shielding system ;
E, temperature-controlled section.

There are two methods, namely, the continuous and the discontinuous method, which can be employed using the above arrangement. With the continuous method, equilibrium is established between the amount of heat entering the apparatus and that leaving it. Let $\theta =$ temperature of the vessel under control; $t =$ ambient temperature; $C(\theta - t)$ be the power emitted by the vessel (conduction, convection and radiation), where C is a positive constant; $P =$ heating power provided by lamps and received by the vessel; $k = -dP/dT$ where k is a positive constant. For normal conditions

$$C(\theta - t) = P$$

and

$$\frac{dP}{d\theta} = C\left(1 - \frac{dt}{d\theta}\right),$$

$$\frac{d\theta}{dt} = \frac{C}{k+C}.$$

Thus the temperature fluctuations in the apparatus depend on those in the room and on the constants C and k. To reduce $d\theta/dt$ so that temperature control within a narrow range of temperature is obtained, it is necessary to make k large compared with C, but under all conditions there will be a change in apparatus temperature with change of ambient temperature.

Using the discontinuous method, the temperature at the surface of the vessel is allowed to fluctuate rapidly from $\theta+\Delta\theta$ to $\theta-\Delta\theta$ and the value of $\Delta\theta$ may be made practically constant by reducing the time-lag between heater and control. With the lag small, the temperature variations at C will be similar to those shown in figure 2 *b*, but if the lag is large, ***hunting*** is said to occur and the temperature variations increase as in figure 2 *c*. In the methods discussed later, the temperature fluctuations are made to approximate to figure 2 *b*, and the value of $\Delta\theta$ is of the order of 0·002° c. Provided that there is some way of removing these fluctuations, the discontinuous method will be more suitable than the continuous one. The variations of temperature at C are conducted into the shield D. If we first consider this to be a semi-infinite solid whose surface temperature is given by $\theta=A\cos(\omega t-\epsilon)$ and whose initial temperature is $\theta=f(x)$, it can be shown (Carslaw, 1922) that if sufficient time has passed to allow the distribution of temperature to become purely periodic, and the influence of the initial distribution has passed away, the temperature of the material is given by

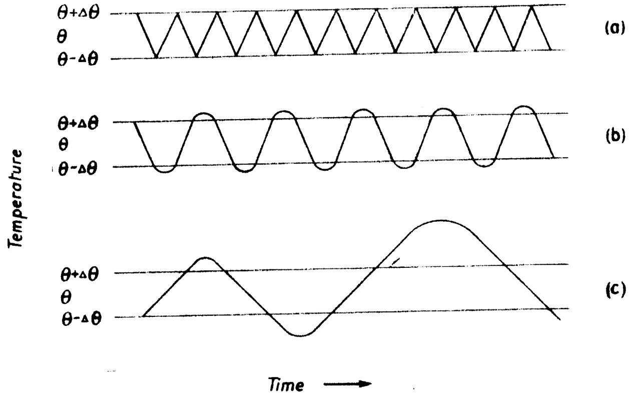


Figure 2. Graphs showing temperature measured by thermometer C for (*a*) no time-lag ; (*b*) small time-lag ; (*c*) large time-lag in control.

$$\theta=Ae^{-\sqrt{(\omega/2k)}}\cos\left(\omega t-\sqrt{(\omega/2k)}\,x-\epsilon\right),$$

where k is the thermal diffusivity of the substance $=K/c\rho$, K being its conductivity, c its specific heat and ρ its density. This shows that a temperature wave is propagated with unaltered period inwards, the amplitude of the wave of shorter period diminishing more rapidly than that of waves of greater period and also having a more rapid alteration in phase. The velocity of propagation is inversely proportional to the square root of the periodic time of the oscillation.

The table gives, for copper and cork, the maximum amplitude and the temperature variation at different distances from the surface when the period of variation of temperature is 1 second and 60 seconds.

Table. θ_x = maximum temperature change x cm. from surface

Substance	k	Period (sec.)	θ_x when x is 1 mm.	5 mm.	1 cm.
Copper	1·13	1	·86	·44	·19
		60	·98	·89	·81
Cork	0·0017	1	·014	10^{-10}	10^{-15}
		60	·57	·06	4×10^{-3}

Rapid elimination of the temperature variations will occur if k is small, x is large and the frequency of the oscillations is high. It is seen from the above table that the temperature fluctuations produced by discontinuous control of a few seconds period can be reduced to a negligible amount by using a layer of cork insulation.

Uniformity of temperature at all points inside E will be obtained if the temperature is the same at all points on the surface C. This can be attained by

(1) arranging the heaters A to produce a uniform heating effect at all points of the surface C;

(2) covering the outer surface of D with a material of high diffusivity such as copper, so that any variations of temperature are rapidly distributed; and

(3) covering the inside surface of D with a material of high diffusivity so that any remaining local variations are eliminated.

The above principles were adopted in the design of the thermostats used for this investigation.

§ 2. METHODS EMPLOYED

Two methods of control have been studied, one employing radiant heat and the other a stream of heated air, and it has been found that, provided the above conditions were satisfied, satisfactory results were obtained by either method. Essentially the methods used consist in making C, the control thermometer, one arm of an alternating- or direct-current bridge. The temperature θ of D is kept a few degrees above ambient temperature and the circuit is arranged so that a drop in temperature $\Delta\theta$ (of the order of 0·002° c.) of the thermometer arm at C will switch on the heaters A. These produce a heating effect which will be removed when the temperature of C reaches $\theta + \Delta\theta$. The time of heating of C can be made approximately equal to its time of cooling. If radiant heat is employed, the time lag of the circuit depends on

(1) the time taken for A to reach a temperature sufficiently high to emit radiant heat in quantity;

(2) the time constant of the electrical circuit.

The first factor is reduced to a very small fraction of a second if electric lamps are used, but these must be uniformly distributed about C. The second factor can also be made very small by using the electrical methods described later. The final time-lag can be made a small fraction of a second.

In the warm-air method, A consists of electric heaters (wire-grid type) behind which are placed electric fans which force air past the heaters and distribute it uniformly around the apparatus. The time-lag using this method can be made a fraction of a second by using fine-wire grid heaters and efficient fans.

§ 3. ELECTRICAL CIRCUITS

(1) *Direct-current bridge method.* The circuit for this method is shown in figure 3. The resistance thermometer is a single-layer coil of No. 44 S.W.G. silk-covered copper of 200 ohms resistance wound round and in close contact

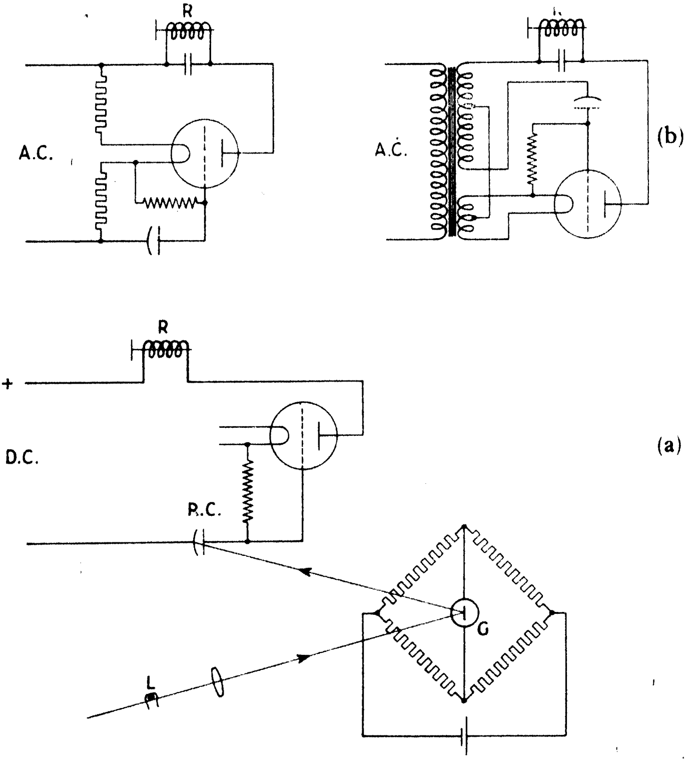


Figure 3. (a) D.C. bridge and photoelectric cell circuits.
(b) arrangements for A.C. supply to valve.
G, galvanometer mirror ; P.C., photoelectric cell ; R, relay ; L, galvanometer lamp.

with part of the apparatus, the temperature of which is to be controlled. This resistance is connected to an equal arm bridge, the other arms being manganin coils. A single accumulator and a reflecting galvanometer are used. The galvanometer should have a stable zero, but it need not be very sensitive. A type found quite suitable was a Tinsley 576-ohm galvanometer which gave a deflection of 150 mm. per microampere. This gave a deflection of 1 mm. for 1/1000° change in temperature. A 921 R.C.A. photoelectric cell and a 1000-ohm mercury relay were used in the circuit. The voltage supply to the valve can be supplied by direct or alternating voltage. Two simple arrangements for using the latter supply are shown in figure 3. In both cases the maximum anode current is one half that which passes when direct voltages are used, as current only passes when the plate is positive with respect to the filament. The circuits shown in figure 3 were found to be very convenient.

Attached to the relay are two or more lamps which are connected to the electric light supply. The lamps are placed a few feet from the apparatus so that they illuminate it uniformly. Efficient shielding will correct for any local irregularities of temperature at the surface.

The variable resistance of the bridge is arranged so that, when the bridge is balanced, the temperature of the thermometer will be three or four degrees above that of the room. The temperature of the external surface rises and falls by a small amount $\pm\Delta\theta$ about the fixed temperature θ. With different instruments this varied from 0·002° c. to 0·005° c. The reduction of $\Delta\theta$ below 0·005° c. is unnecessary, as this small fluctuation is rapidly diminished as it penetrates a small distance into the surface of the apparatus. This method was used when radiant heat was used as the heating source.

(2) *Alternating-current bridge method.* A discontinuous method of control was devised, using an alternating bridge with valve amplification. The circuit used was similar to that given by Sturtevant (1938), the chief modification being that the high-power thyratron which gave continuous control was replaced by a power valve or low-power thyratron and relay to give discontinuous control. This simplifies the operation of the circuit, gives a more suitable type of control, and greater sensitivity.

One arm of the bridge consists of the thermometer (190 ohms at 15° c., 44-gauge copper resistance, non-inductively wound) and a variable manganin resistance of 10 ohms, the other arms being each of 200 ohms manganin wire, non-inductively wound. To obtain temperature control, this variable resistance is reduced so that for bridge balance the thermometer temperature must rise a few degrees above room temperature. Maximum current at first passes through the relay and the lamps or heaters are switched on. As the temperature of the apparatus rises, bridge balance is approached and the current through the relay falls until the heating supply is suddenly switched off. The lamps or heaters then operate every two or three seconds. It was found that a change of temperature of about ±0·003° c. in the thermometer arm of the bridge caused a change of 1 ma. in the output current, and this is sufficient to operate the relay.

Comparison of the bridge methods. For the radiant-heat method, both bridges were found to be equally good, although the A.C. bridge circuit was more elaborate. A very compact D.C. bridge arrangement can be made using a box galvanometer and inserting the photoelectric cell in place of the scale. With the heated-air method, the fluctuations in temperature $\Delta\theta$ are not exactly equal and, using a sensitive galvanometer of long period, the spot might overshoot the photoelectric cell and so produce additional heating without control. An electrical time-switch or some such device must be added to prevent this lack of control causing any damage. With the A.C. bridge method this does not occur, as an increase in temperature of the thermometer arm reduces the current through the relay.

§ 4. CONCLUSIONS

Where efficient thermal shielding can be employed, the radiant-heat method, owing to the rapidity with which heat is transferred from the heater to the apparatus, is sounder theoretically and practically. The main difficulty in this method is in the production of uniformity of temperature in the apparatus, and the methods for obtaining this have been discussed earlier. If control is required over a large volume and little or no thermal shielding is available, the heated-air method is superior, as it produces a more uniform heating effect. For precision results, using this method, it is advisable to protect the apparatus from all sources of radiant heat; windows and lamps, etc., should thus be covered. The thermometer may be placed close to or in contact with the apparatus whose temperature is to be controlled, and the observer should not stand near the apparatus.

The radiant-heat method was used in the experimental determination of the electronic charge, and the temperature inside the apparatus was measured by means of a sensitive fine-wire thermocouple. No temperature variations were detected throughout the tests, the temperature being read to 0·01° C. Using the heated-air method, partially or entirely uninsulated large vessels were temperature-controlled to better than 0·01° C., the temperature for these tests being read by means of a sensitive mercury-in-glass thermometer.

§ 5. ACKNOWLEDGEMENTS

This investigation was commenced during the latter part of 1938 at the Natural Philosophy Laboratory, University of Melbourne, and was continued during the following year at the Physics Laboratory, University of Western Australia. The writer wishes to acknowledge the facilities he has been given at both these universities. The investigation was made possible by the help of a Commonwealth Research Grant.

REFERENCES

Laby, T. H. and Hopper, V. D., 1939. *Nature, Lond.*, **143**, 240.
Callender, A., Hartree, D. R. and Porter, A., 1935-6. *Philos. Trans.* **235**, 415.
Carslaw, H. S., 1922. *Mathematical Theory of the Conduction of Heat in Solids* (Macmillan & Co., Ltd.), p. 47, second edition.
Sturtevant, J., 1938. *Rev. Sci. Instrum.* **9**, 276.
Turner, L. B., 1937. *J. Instn. Electr. Engrs*, **81**, 404.

AN APPROXIMATE SOLUTION FOR THE DISTRIBUTION OF TEMPERATURE OR POTENTIAL WITH CYLINDRICAL ISOTHERMAL OR EQUIPOTENTIAL SURFACES

By S. WHITEHEAD,
British Electrical and Allied Industries Research Association

Received 20 *November* 1941

ABSTRACT. Dipoles or source-sink combinations are used to give an approximate solution of Laplace's equation when the boundary surfaces comprise equipotential or isothermal cylinders. The method is shown to be convenient for a symmetrical arrangement but liable to become laborious in other circumstances.

The method of solution for the field between two equipotential or isothermal spheres is well known and has been fully analysed by Russell (1911). If the centre of sphere 1 is O_1 and that of sphere 2 is O_2, then charges or sources are placed at O_1 and O_2, as required, to bring spheres 1 and 2 considered separately to their requisite potentials or temperatures. But the charge or source at O_1 causes sphere 2 to have a variable potential or temperature; accordingly, O_1 is compensated by its image (at the inverse point) in sphere 2 and O_2 by its image in sphere 1. This leaves the images uncompensated, which may be rectified by the 2nd-order images and so forth. Since further, the magnitude of the image decreases with its order, a convergent series is obtained and the problem has an exact solution. In the same manner, but with much greater labour and with additional restrictions, a solution could be obtained for more than two spheres. But in two dimensions, with spheres replaced by cylinders, the images are equal and alternatively positive and negative, so that the series is no longer convergent but oscillates, and is also divergent if there are more than two cylinders or a boundary condition other than at infinity, because the number of images of a given order then increases with the order.

If a solution were possible by means of a series of positive and negative point-charges or sources and sinks, then the resultant image-charge or resultant source would be zero. This suggests the use of a dipole or source-sink combination as a method of approximation, a method which has been found in fact to give a reasonable closeness of fit without undue numerical labour. Mathematically, such a method is valid since a solution of Laplace's equation may be built up from a distribution of point charges and dipoles for any space bounded in such a way as to exclude the charges and dipoles.

An example will make the method clearer. Consider three similar cylindrical isothermal or equipotential surfaces in a semi-infinite medium within which heat is generated at a constant rate of $g/2\pi$ watts per cm. length or which are charged with a uniform electrification of $\epsilon/2\pi$ coulombs per cm. length, where g is the thermal resistivity (in watt-cm.-°C. units), or ϵ the permittivity of the medium. There are many practical examples of this problem, as with three charged conductors near an earthed plane or three underground pipes.

Let the centre of the system be C; let the centres of the individual cylinders be O, O′, O″ respectively. The figure shows one of the cylinders considered. "Unit" sources of heat at O, O′ and O″ give the correct rate of heat generation, but the circles are not isothermal. If this may be corrected by dipoles, which do not alter the rate of heat generation external to them, these dipoles must by symmetry be situated on CO, CO′, CO″; for example, at D, D′ and D″. If the strength of the dipole is M, the temperature it produces at any point P is Mx/r^2,

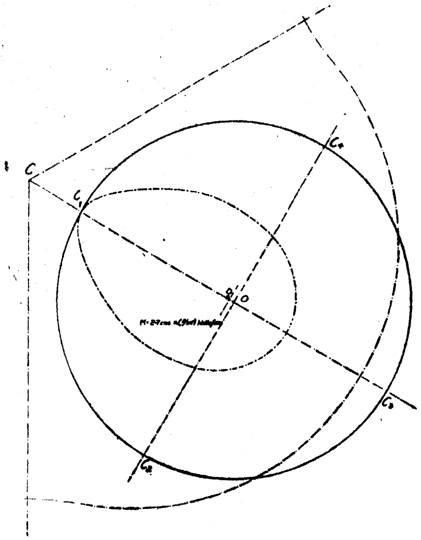


Figure 1.

Solution for OC=5·08 cm., OC=6·8 cm., CD=6·65 cm. and M=2·7 cm. ×(g/2π) watt/cm.

where r is the distance between P and M, while x is the projection of r on the direction of M, and is positive if measured in the same direction as M. In the case mentioned, the directions of the dipoles must be DO or OD, etc. To make the plane surface an isothermal, the images of O, O′ and O″ in the plane surface are occupied by "unit" sinks. To simplify the problem, the distance of the plane is supposed to be sufficiently great for the variation of temperature over the circles due to these images to be neglected, together with the variation of temperature over the plane surface due to the dipoles.

The position and magnitude of the dipoles may then be determined by applying the condition that the temperatures at C_1, C_2, C_3 are equal, which will then also be true of C_4. Whence, if h is the depth of C below ground surface:—

$$\log\frac{2h}{O'C_1}+\log\frac{2h}{O''C_1}+M\left[\frac{DC_1}{(DC_1)^2}+\frac{D'C_1'}{(D'C_1)^2}+\frac{D''C_1''}{(D''C_1)^2}\right]$$

$$= \log\frac{2h}{\mathrm{O'C_2}} + \log\frac{2h}{\mathrm{O''C_2}} - M\left[\frac{\mathrm{DO'}}{(\mathrm{DC_2})^2} - \frac{\mathrm{D'C_2}}{(\mathrm{D'C_2})^2} - \frac{\mathrm{D''C_2''}}{(\mathrm{D''C_2})^2}\right]$$

$$= \log\frac{2h}{\mathrm{O'C_3}} + \log\frac{2h}{\mathrm{O''C_3}} - M\left[\frac{\mathrm{DC_3}}{(\mathrm{DC_3})^2} - \frac{\mathrm{D'C_3}}{(\mathrm{D'C_3})^2} - \frac{\mathrm{D''C_3''}}{(\mathrm{D''C_3})^2}\right],$$

where $\mathrm{C_1'}$, $\mathrm{C_1''}$, etc., are the projections of $\mathrm{C_1}$ on D′C, D″C respectively.

Thus $$\log\frac{\mathrm{O'C_2\,O''C_2}}{(\mathrm{O'C_1})^2} = (K_2-K_1)M$$

and $$2\log\frac{\mathrm{O'C_3'}}{\mathrm{O'C_1}} = (K_3-K_1)M,$$

where K_1, etc., $$= \frac{\mathrm{DC_1}}{(\mathrm{D'C_1})^2} + \frac{\mathrm{D'C_1'}}{(\mathrm{D'C_1})^2} + \frac{\mathrm{D''C_1''}}{\mathrm{D''C_1}^2}, \text{ etc.}$$

The geometry of the problem then permits the determination of M as $af(a/b)$ and CD as $a\,\phi(a/b)$, where a is the radius of the circles, b is OC, and f and ϕ are functions determinable from the equations given. The algebraic expression of these functions is not given here, since it is usually more convenient to measure on a diagram and to solve the equations graphically. The figure shows the solution of these equations for $a=5\cdot08$ cm. (2″), $b=6\cdot8$ cm., $h=88\cdot1$ cm., giving $M=0\cdot53a$, $\mathrm{CD}=1\cdot31a$, i.e. $f(0\cdot747)=0\cdot53$, and $\phi(0\cdot747)=1\cdot31$. The broken lines show two isothermals for the case when the dipoles M are omitted; one passes through $\mathrm{C_1}$, while the second corresponds to the temperature attained by $\mathrm{C_1}$ when the dipoles are included. The isothermal corresponding to the temperature of $\mathrm{C_1}$ (and $\mathrm{C_2}$ and $\mathrm{C_3}$) when the dipoles M are included coincides so closely with the circle that it cannot be shown separately on the scale employed. The actual value of the temperature calculated for various points on the circle does not depart by more than 0·4 per cent from the value at $\mathrm{C_1}$, $\mathrm{C_2}$, $\mathrm{C_3}$.

A solution can similarly be obtained for any symmetrical disposition of circles, since the dipoles are located on the lines joining the centre of the complete figure with the centres of the circles. With a completely unsymmetrical arrangement, the method is in many instances too complicated to be worth while, since within each circle is a dipole corresponding to each other circle and situate on the line of centres. For a row of circles, this however reduces to two dipoles in each circle situated symmetrically with respect to the centre on the line of centres, except for the two outer circles, where the outer dipoles are missing.

When the circles are small compared with their axial spacings or distance from the plane surface, the greater accuracy obtained by the dipole method is unnecessary and the principle of superposition may be followed as in a specific example, computed by the present writer (1930).

REFERENCES

Russell, A., 1911. *Proc. Phys. Soc.* **23**, 352; 1912. *Ibid.* **24**, 22; 1914. *Alternating Currents*, vol. **1**, 2nd edition. (Cambridge.)

Whitehead, S., 1930. *World Power*, **13**, 350.

A ROTATING DIFFERENTIAL PHOTO-ELECTRIC PHOTOMETER FOR PRECISION WORK, by J. T. MacGregor-Morris and A. G. Stainsby *

DISCUSSION

Dr. J. R. **Tillman**. I was surprised that no mention was made either of methods using A.C. amplifiers or of attempts to do the commutation with thermionic valves. If generous use is made of negative feedback, amplifiers having an extremely high stability against variations of power supplies and giving a very faithful reproduction of wave-form can be made very cheaply. Thus, if two stationary photocells were used, half a cycle out of phase with each other, and receiving light alternately from the variable source and the constant source via, say, a rotating double-sided mirror, each being followed by a good amplifier, the resultant signals could be placed in opposition and the difference, if any, further amplified. A robust meter could now be used. Small differences in the sensitivities of the two cells could be compensated by using a gain control in one amplifier. It might be possible to use only one photocell if the wave-form of the signal obtained could be accurately analysed. The amount of some of the higher harmonics present will change, as the alternate signals (from the variable source and the constant source) vary in relative intensity, by much larger percentages than does the intensity of the fundamental frequency.

I have recently tested a batch of six commercial " rectifier " cells of high sensitivity (480 μa./lumen) and have found good consistency. Thus four of the cells gave sensitivities within 1 per cent of the mean; other properties might of course differ by larger amounts. None the less, I think stress should be laid on one of the inferences to be drawn from Prof. Morris's paper, namely, that it is very difficult to measure, rapidly, small changes in the intensity of a light source, even when use is made of good " rectifier " cells.

Mr. J. S. **Preston**. Professor Morris has described a method of photometry with rectifier photocells which not only eliminates electrical troubles, but also disposes artistically of all those features of a cell which the photometrist regards as perverse because he can neither predict the magnitude of their effects nor satisfactorily explain their causes. It is, I think, correct to describe this as a new achievement, at least technically, in view of the great difficulties encountered in past efforts to use a similar principle of alternation with the emission photocell followed by amplifiers. Here the power available was larger, but the degree of success attained considerably less, I believe.

The method remains a balancing method, however. It would be of value to discover the true causes of fatigue, so that cell manufacturers might try to eliminate them, and so enable the rectifier cell to be used with greater precision as a direct measuring instrument, since in other respects it is so admirably suited to this method of use.

---

* See *Proc. Phys. Soc.* **53**, 584.

Professor Morris mentioned the effect of varying the angle of incidence of the light falling upon a cell. I do not know whether the "obliquity characteristic" has any important effect on the wave-form of the output from his photometer, since the cells rotate. It may be interesting to note that measurements made at the N.P.L. show that all cells coated with a smooth lacquer behave very similarly as regards loss of efficiency with increase in the angle of incidence of the light. This is so irrespective of the appearance, matt or otherwise, of the selenium.

If the amount or flux of light falling on the cell surface is maintained constant as the angle of incidence θ is varied, then the current output of the cell varies as F_θ, Fresnel's value for the fraction of the incident light which is transmitted by a single refracting surface, the appropriate value of the refractive index being about 1·55. Or, of course, if the cell is at a constant distance from a constant light source, and is rotated, the output current varies as $\cos\theta \, . \, F_\theta$. (In both cases the illumination-current characteristic of the cell is assumed here to be linear.) This suggests that the cell is equally efficient for all the light which actually enters the first lacquer surface, irrespective of the angle of incidence. The second interface, namely, lacquer-sputtered metal film, appears to be without influence in this respect, but it may be noted that the light incident on this interface is confined within a cone of semi-vertical angle equal to the critical angle for a lacquer-air surface.

Mr. L. T. MINCHIN. It may be of interest to mention another name frequently used for these cells*—" blocking-layer ". I presume both this and " barrier-layer " are translations of the German name " Sperrschicht ".

Do I understand that pairs of commercial cells are not suitable for this method? Is it necessary to have them specially made at the N.P.L.?

Mr. D. W. STOPS. I should like to ask Professor Morris how it was ensured that the two photocells employed in his photometric instrument gave equal e.m.f.s for equal sources of light and for equal periods of time.

Lt.-Col. K. EDGCUMBE. This paper is of particular interest to photometrists as outlining a method which would seem to combine the well-known advantages of objective photometry with an accuracy even greater than that of the visual. The most serious limitations, although not stated by the authors, would appear to be those due to the colour response of the cells themselves. It is likely, however, that when colour filters are used, a precision comparable with that of the visual method might still be possible.

---

* In presenting the paper, Prof. Morris mentioned that the names *Vorderwand* and *barrier-layer* are both in use for this type of cell. [ED.]

The authors stress the point that their chief objective was to obtain relative readings with a rapidly varying source, viz., an electric arc. If this was all, it would appear that a photo-electric cell connected to a galvanometer would have given the information required and would have enabled a curve such as that shown in figure 9 to be drawn with considerably less complication. It appears to me that the chief field of usefulness for this method will be the comparison of two similar light sources to a high degree of precision.

Turning to the method itself, my own feeling would be that the advantages of using a single cell would far outweigh the difficulties, and as regards the trouble experienced, due to the heavy alternating current, it would seem that it should be possible to damp this out by the use of a high-capacity condenser. In any case, I think the authors should use the current output through a low resistance as the criterion, rather than the voltage. They appear to assume that the only advantage of the current method lies in its increased sensitivity, but they would find that many of the troubles, particularly those due to temperature and drift, are considerably reduced and that almost perfect linearity is possible with a moderate galvanometer resistance. In this connection, the bending over of the curve of figure 6 seems to show that the cells are being used under adverse conditions.

The differentially wound galvanometer discarded by the authors would seem to be the simplest way of applying the current criterion; but if this is undesirable, then the joining of the two cells in series, with the galvanometer so connected that it carries the difference current only, should not be difficult. As commutation would always occur during the period of zero illumination, the fact that the circuit was periodically opened should have no effect upon the performance of the cells.

This latter consideration leads me to ask whether the authors have found that the characteristics of the cells are, in general, the same when subjected to a rapidly varying illumination, as they are with a steady illumination of the same average value.

AUTHORS' reply (by Prof. MACGREGOR-MORRIS). Mr. Tillman has described a neat way of doing away with the use of a commutator, and in the hands of one experienced in the precision use of such apparatus, it is practically certain to give as good results as ours. As to the question of relative costs, I am doubtful ; and I naturally use devices with which I know the traps and how to overcome them. Other experimentalists more conversant with the use of amplifiers and similar apparatus—though they have their defects—might well get as good results, and that without any noise or vibration.

Mr. Preston's encouraging remarks will certainly help us to persevere. My only doubt is whether he has found such similarity in behaviour that the loss of efficiency with the increase in the angle of incidence of the light is unaffected.

In reply to Mr. Minchin, as these were the first cells to be used for this purpose, and since I had friends there, I decided to go the N.P.L. in the first place, and the cells have turned out exceptionally. Now, however, that more experience has been gained by manufacturers, I would certainly turn to them.

In reply to Mr. Stops, the method is given in the paper, and it has to be remembered that our apparatus turned at 10 revs. per second, and so the length of time is very short.

Lt.-Col. Edgcumbe's remarks are very much to the point. I fully agree that there are other fields where this instrument will have much wider application. But the simple arrangement outlined by Col. Edgcumbe was used by us until the repeated determination of all the corrections became so tedious that for the sake of time we were forced to develop the method described in the paper, and it has certainly saved much time and has added to the accuracy of our results. Turning to his next point, the principle involved in using a large-capacity condenser (if it is perfect) is sound, but my experience of the lack of perfection, unless a very good one is used, made me decide to put this method on one side. It is possible that I am out of date in this matter.

The cells were actually used at 10 f.c. usually, and it is the bottom graph in figure 6 which applies (i.e. the Barnard cells), where the graph is not far from straight.

REVIEWS OF BOOKS

The Second Yearbook of Research and Statistical Methodology Books and Reviews, edited by OSCAR KRISEN BUROS. Pp. xx + 383. (The Gryphon Press, Highland Park, New Jersey, 1941.) $5 (less 10 % if ordered direct from the publishers.)

Dr. Buros is disturbed, and rightly so, because text-books dealing with statistical and allied subjects often betray the ignorance of their authors in the recent developments of that subject. He argues that if the writers are ignorant, those teachers who do not publish are likely to be in even worse case ; and if the teacher is behind the times, what about the pupil ? To ameliorate this position, Dr. Buros had the idea of collecting together the titles of books on the subjects concerned, and printing a collection of reviews of each book. Among the purposes so served are (1) that those conscious of their defects could find the names of books for study, and could receive guidance as to which were good and which bad ; (2) presumably the bad books would be pilloried ; (3) the fact that some reviewers are woefully ignorant of the subject, or else too kind or too polite to give frank reviews, would be exposed. That this happens is, I think, undoubted. That it is worth a whole book to illustrate it is very doubtful.

So much for the project. Next, for its accomplishment. In 1938, Dr. Buros brought out a work, covering, if I have understood him correctly, the books reviewed between 1933 and 1938. The second yearbook, the one now under review, deals with 359 books published from 1933 onwards and contains 1652 reviews or parts of reviews of them. The prospective purchaser will naturally ask, " What books, and what reviews ? " The author's answer is that in the first edition he had attempted to include " methodology books in all fields ", and he now says " Some classes of methodology books (e.g., books on microscopy and the philosophy of science) have been excluded." Perhaps the first few titles will give as good an idea as any other explanation. They are :—Statistical Mathematics, The Nomogram, Mathematical Analysis for Economists, Family Expenditure, Educational Research, The Technique of Marketing Research, Time-series Charts, Bibliographical Citation in the Social Sciences, Graphs, Statistical Methods, Practical Application of the Punched card method . . ., The Preparation of Reports, A History of Historical Writing, and The Study of Society.

As to the choice of reviews, the author has no criterion of selection by merit. He has a list of 283 journals, and includes excerpts from all reviews of a book which appear in any of these journals. Thus we find Pledge's *Science since 1500* reviewed in *Isis*, the *Journal of the Royal Microscopical Society*, *Nature*, the *New Statesman*, *Philosophical Magazine*, our own *Proceedings*, *Science and Society*, the *Spectator*, *Time and Tide* and the *Times Literary Supplement*. Surely a library could risk 7*s*. 6*d*. without this weight of testimony ? The only reviews of Burington's *Handbook of Mathematical Tables* (the name is wrongly given as Burlington in the heading and in the catch-word at the top of the page, but correctly in the index) are from the *American Mathematical Monthly*, the *Journal of the American Society of Agronomy*, *News Edition of the American Chemical Society*, and *Popular Astronomy*.

Since only excerpts are given from the reviews, the question of their accuracy of representation arises. I find four of my own reviews, and I am only slightly misrepresented in one place. Dealing with Miller's Tables for converting rectangular to polar co-ordinates, I expressed the view that the scheme for ascertaining the quadrant in which the radius vector lies had no advantage over the direct deduction from first principles ; the omission of a few words makes it appear that I thought the *whole* scheme had no advantages over an *ad hoc* calculation in each particular case.

In the introduction, the editor mentions possible improvements in future editions. Of these, the most important seems to be the inclusion of books published in languages other than English, for many of the most important works are necessarily omitted under the present restriction. The inclusion of foreign books will, however, have the effect of swelling the volume considerably, and attention might well be given to the question of defining more clearly the field covered and of pruning vigorously the list of journals from which reviews are taken.

There is also a suggestion that articles, as distinct from books, might also be reviewed in future editions. This plan has much to commend it, though it might be still better to let experts give a more general review of progress, taking into account those papers which they regard as important. No doubt experience will show the best form for these articles to take.

There is an elaborate system of indexing—lists of journals from which reviews have been taken, an index of reviewers and indexes of the books reviewed. I have enjoyed reading the book, or bits of it selected at random, when perhaps I ought to have been doing something more active, but I am not certain that it would be everyone's taste.

J. H. A.

Practical Solution of Torsional Vibration Problems, Vol. 2, by W. KER WILSON. Second Edition. Pp. xx + 694. (London: Chapman and Hall, Ltd., 1941.) 42*s*.

The first six chapters of this very useful work were included in vol. 1, whilst vol. 2 comprises six further chapters, as follows :—Chap. 7, Determination of Stresses due to Torsional Vibration at Resonant Speeds ; Chap. 8, Measurement of Torsional Vibration Amplitudes and Stresses ; Chap. 9, Analysis of Torsiograph Records ; Chap. 10, Torsional Vibration Damping Devices ; Chap. 11, Rotating Pendulum Vibration Absorbers ; Chap. 12, Dynamic Characteristics of Electrical Generating Sets Direct-coupled to Internal Combustion Engines, and Flywheel Calculations for Marine Installations.

Broadly, vol. 2 treats of the subjects of damping, the measurement of vibration and means for the elucidation of these measurements. The most recent developments are included, as, for instance, the strain-gauge technique, which promises to pave the way for rapid progress in the practical measurement of vibration stresses in airscrew blades. There is also a fairly extensive treatment concerning the utilization of rotating-pendulum devices to overcome specific orders of resonant torsional vibration in the running range.

The strain-gauge procedure may be said to be in its infancy, and there is some question as to the accuracy of the results in its present stage. The author gives a description of an electrical method which is grounded on strain-sensitive resistance strips, which strips respond to changes in the surface strain at selected points on the blades. The rotating pendulum " damper " has the remarkable property that it can be made to nullify the effect of a particular order of forced torsional vibration. Care must, however, be taken to ensure that other orders previously innocuous do not become accentuated. As with practically all anti-vibration devices, it alters the original system by modifying its frequencies and modes of vibration and, in addition, each pendulum employed introduces an extra frequency.

As a compendium of the various notions and devices pertaining to torsional vibration which have become crystallized, the author's work may be classed as comprehensive. There is even included in volume 2 a *Selected List of Patents relating to Vibration Study*. In the placing of these devices in their true perspective it must be realized that an artifice which will diminish vibration in one set of circumstances may increase it in other cases, or at other speeds in the same operating range. So far as is possible, vibration characteristics should be taken into account in the design stage. If the trouble and its main cause are clearly understood it may be found that appropriate alterations in the dimensions of a particular shaft, for example, may be all that is called for, whereas the insertion of a so-called vibration damper may have little influence in eradicating the trouble, and may even amplify it.

Certain items dealt with in the book are from their very nature empirical and imperfectly understood, as, for instance, elastic hysteresis and apparent damping in engine crankshafts. Care must be taken to ensure that a particular effect is not over-estimated whilst other equally or more important effects are neglected. For example, it has become customary to calculate the dissipation due to hysteresis, which may be only 10 per cent of the total dissipation, and to multiply it by an empirical factor determined so as to make the calculated torsional amplitude coincide statistically with the measured one in a number of installations (Den Hartog).

Torsional vibration is only one part of the practical story. Resonant lateral vibration can, in certain cases, be more serious than torsional, and frequently the two are indissolubly coupled.

Like the first, the volume under review is copiously illustrated, both diagrammatically and photographically. J. M.

Visibility in Meteorology, by W. E. KNOWLES MIDDLETON. Second Edition. Pp. x + 165. (Toronto, Canada: The University of Toronto Press; London: Sir Humphrey Milford at the Oxford University Press, 1941.) $2.50.

The appearance after six years of a second edition of this monograph is a sign that Mr. Middleton has supplied a want and that he has done so in a not unsatisfactory fashion. The author has, in fact, produced a stimulating account of researches and meteorological practices bearing, as he himself puts it, on the question, " How far can I see to-day ? " Despite certain shortcomings of the work, all interested in the visual range of objects and lights by day or night will find Mr. Middleton's book of value.

Compared with the first, the second edition contains additional matter on the optical and other properties of aerosols and on the perception of lights and brightness contrasts under different conditions. The chapter on instruments is amplified and a much extended discussion is given of practical methods of observing visual range and of recording and analysing the results. The forecasting of fog forms the subject of a new chapter. An adequate theory (due to the author) of the visual range of coloured objects is included as an appendix and the bibliography is increased from 146 to 342 references.

The theory of the air-light bulks largely, and the author rightly explains in some detail Koschmieder's rather complex analysis. It is a pity, however, that the simplest proof of Koschmieder's main result has been overlooked. The apparent brightness H_s of a small black object consists simply of the air-light of the layer of atmosphere between observer and object. If the object is removed the observer sees, added to the air-light, the brightness H_h of the sky beyond the object, reduced by the transmission factor $e^{-\sigma l}$ of the intervening atmosphere. Thus, provided the sky brightness does not change as the observer moves up to the position of the object, keeping his direction of view the same, we have $H_h = H_s + H_h e^{-\sigma l}$, which is Koschmieder's result.

If the object is not small, then on its removal the air-light will be increased because rays which previously were obstructed by the object will reach and be scattered in the intervening atmosphere. On this account, Koschmeider's expression gives too large a value for the apparent brightness of the black object. Mr. Middleton refers briefly to Foitzik's work on this question but gives more attention to the complication which ensues when the object is so small that diffraction of light round the edges may play a rôle. Löhle's theory of this edge effect is certainly not easy to follow, but Mr. Middleton reproduces it uncritically and derives a formula for the apparent brightness at the centre of a black object, which involves explicitly neither the distance of the object nor the coefficient of extinction. The uninitiated reader is left puzzling how Löhle's formula is linked with Koschmieder's result. This is a good example of the main defect of the work, the author's tendency to present condensed accounts of theoretical or experimental researches without adequate welding of these members into a single structure. Some needless repetition, as well as lack of continuity, are the inevitable consequences of this tendency.

Of the few minor errors and obscurities, two will be mentioned. The table on p. 90 includes values for the visual range of objects seen against the earth's surface, computed from formula 5.16. This formula, however, applies only when object *and* background present vertical surfaces. When the background is horizontal, the formula must be modified by the substitution of $2R_2$ for R_2. On p. 48, Holladay's glare formula is quoted with the numerical value 10 for the constant K. It should have been stated that the units implied are candle/sq. ft. for brightness, foot-candle for illumination and degree for angle; the first two of these are not the units commonly employed in the rest of the book.

The author is at his best when giving his own views, as in the appendix, and when he is not following his authorities too closely. He might bear this point in mind when preparing the further edition which in due time will surely be required. W. S. S.

Practical Physical Chemistry, by ALEXANDER FINDLAY. 7th Edition. Pp. x + 335. (London: Longmans, Green and Co., Ltd., 1941.) 12*s*. 6*d*.

In such a large and ever expanding field as that of practical physical chemistry, no single book can satisfy the likes and dislikes of every individual reader. It is, however, probably true to say that Professor Findlay's book meets the varied demands of the subject more successfully than any other English text.

The present edition follows closely that adopted in earlier ones, and it is a tribute to the original planning that the new matter has fallen into place without destruction of the logical sequence. The book has, since the last edition in 1935, been revised and enlarged; the use of small type in certain places has enabled this to be carried out without greatly increasing the bulk. The production of the book has vastly improved, and it has become correspondingly more readable. The use of bold type for the general constants is to be commended in that it familiarizes the student at the start with the present system of scientific nomenclature.

New material in the present edition includes: the use of the thyratron relay, the vapour-density methods of Weiser and of MacInnes and Kreiling, the maximum bubble-pressure method of measuring the surface tension of liquids, the use of the antimony electrode in determinations of hydrogen-ion concentration and of the tungsten electrode in redox systems, boiling-point curves of binary liquid mixtures and equilibrium curves in three-component systems. The number of experiments and of applications of experimental methods described has been increased. References to the literature have largely been brought up to date, as has some of the text itself—for example, the discussion of surface tension and of viscosity in relation to constitution.

A few of the chapters still remain rather unsatisfactory. The treatment of colloids, of surface phenomena and of reaction velocity are lacking both in appeal and in completeness. Also there is no direct experimental study of the kinetic theory. It is to be hoped that a later edition of this deservedly successful book will remedy these deficiencies. R. F. B.

Diffusion in and through Solids, by RICHARD M. BARRER. Pp. xiii + 464. (Cambridge: The University Press, 1941.) 30*s*.

A great deal of the progress in science is due to the painstaking work of scientists who undertake the task of reviewing, systematizing and tabulating the results of experiments. A book on a subject not hitherto dealt with in such detail is sure of a welcome.

As the measurement of permeability and diffusion constants forms the main substance of this book, the author gives details of the solution of the diffusion equation needed in the discussion of typical experiments. The flow of gases through narrow channels is also discussed, with its application to flow through porous materials. A series of chapters are devoted to the absorption of gases in, and diffusion through, metals and alloys, including the very interesting hydrogen-palladium and deuterium-palladium systems. Less well-known aspects also dealt with include the permeability to gases and vapours of complex silicates, and of complex organic substances—rubber, resins, leathers, paper, etc. Inter-diffusion and self-diffusion in the case of metals, and ionic diffusion in the case of alkali and silver halides, form the subject matter of special sections. A brief account is also given of experimental and theoretical work on the mobility of atoms on surfaces.

For a proper appreciation of the processes of diffusion it is necessary to marshal a large array of facts from many branches of physics. The structure of complex organic and inorganic substances, the nature of the imperfections in otherwise "perfect" crystals and facts about F-centres and U-centres (in ionic crystals) form examples of such subsidiary matter.

The theoretical side is, in general, adequately summarized and the documentation is very complete. A special feature is the collection into tables of a great mass of experimental data, which should prove of great value to physicists and chemists alike. M.B.

The Birth and Death of the Sun, by GEORGE GAMOW. Pp. xiv + 238. (London: Macmillan and Co., Ltd., 1941.) 12*s*. 6*d*. net.

This excellent book gives a very clear and interesting account of the latest views of the relevance of nuclear physics to the problems of stellar constitution and evolution. At last a plausible idea of the mechanism by which stellar energy is made available for radiation can be formed, and Professor Gamow here explains it in a form suitable for the ordinary non-scientific reader. A noteworthy feature of the presentation is the provision of many original and ingenious diagrams, which give a very vivid idea of the rather abstract processes described. From the layman's point of view the book is slightly marred by a tendency to present useful hypotheses as established facts, and a somewhat excessive optimism (or should we call it pessimism ?) which leads the author to declare, for example, that " it may be hoped that in the course of the next few years a satisfactory solution of this last remaining puzzle of stellar evolution will finally be found ". Physicists of experience will not attach too much weight to this, and the book can be recommended to them almost unreservedly as the best description we know of the present state of the matters discussed. This having been said, it will not be out of place to point out a few defects, with the remark that their importance is negligible compared with that of the merits of the book.

It is not true that " it was first indicated by the Russian chemist, Dmitri Mendelyeev " (p. 48) that the elements show a periodicity in properties. Newlands indicated this before Mendelyeev, but his work met with a cold reception. Plate VII (p. 126) is the worst of many bad illustrations which exist of the Harvard sequence of stellar spectra. Not only is it not stated that most of the spectra shown are of the iron arc and not of stars, but the stellar spectra themselves show next to nothing of the characteristic features of their types. Capella is not a typical red giant (p. 141), nor do " the bright emission lines of novae show a conspicuous shift toward the violet end of the spectrum " (p. 185). They are broadened but undisplaced, and the absorption lines at their edges show the shift. The disc-like form of the stellar system was, in a sense, " proposed by Herschel more than a century ago " (p. 207), but it was clearly described by Thomas Wright, of Durham, nearly two centuries ago. Finally, we know of no evidence that the two nebulae apparently near the great Andromeda nebula are " satellites " of the latter (p. 218).

H. D.

Weather Analysis and Forecasting. A Textbook on Synoptic Meteorology, by SVERRE PETTERSSEN. Pp. xvi + 505. (London: McGraw-Hill Publishing Co., Ltd., 1940.) 35*s*.

The concepts *air mass*, *front* and *wave cyclone* to-day form the framework of the weather forecaster's mental processes. They represent, so to say, the leitmotifs of the weather tone poem, give it line and form where the untrained ear finds only notes strung together rather incoherently. As yet, the structure of the piece is by no means thoroughly understood, but at least the stage has been reached when it is worth while for the critic's analysis to be given, and, very fittingly, one from the Norwegian school of meteorologists, who have been mainly responsible for the development of the concepts and their interplay, has now undertaken the first thoroughgoing presentation of the analysis. For thoroughgoing this book attempts to be, and is not to be compared with a number of introductory treatments of synoptic meteorology—the meteorology of the weather map—already provided. It seeks to present a unified picture of the laws of dynamics and thermodynamics working themselves out on the macroscopic scale in the weather sequence, day by day. Full mathematical treatment is accorded where meteorological thought is sufficiently clarified to allow this ; some readers indeed will feel that the mathematical horse has been ridden too hard, not because of any difficulty in the mathematics, but because its results often seem of doubtful validity in a subject in which to know what to neglect and what to consider is half the battle.

The book opens refreshingly by a chapter headed *Air Mass Characteristics*, followed by *Stability and Instability in Relation to Weather Phenomena*. These 137 pages show the author at his best, though many meteorologists will not assent to all that is in them. They are pointed by excellently conceived diagrams—a feature indeed of the whole book—giving succinct presentation of the physics of a multitude of weather phenomena, and more than mere illustration. A more pedestrian but necessary chapter on the production and transformation of air masses follows, and then we pass to the kinematic aspects of weather: wind and the formation of fronts, with a discussion of the latter's characteristics. The kernel of the book is then reached with 50 pages devoted to waves and cyclones, and here least is understood, so that a completely successful treatment is not at present possible. Some integrating physical idea is still awaited, and that in turn may wait on more data being acquired in a more purposeful way.

The remaining four chapters deal with the analysis of weather charts and the methods of weather forecasting. Surface and upper-air charts and the vertical cross-section are each considered, and Namias, an American meteorologist, contributes one of these chapters, entitled *Isentropic Analysis*, showing how a suggestion made by Sir Napier Shaw many years ago is beginning to show some fruit. Petterssen's personally devised kinematic method is not perhaps worth all the space devoted to it—meteorological data and the forecaster's time are rarely sufficient to enable second and third order differentials of pressure with respect to time and position to be evaluated. But the last chapter, giving examples of charts analysed, has its feet on the ground, and the value of many of the 31 concisely expressed rules, previously derived, is well demonstrated.

In this book much has been attempted and much done. No meteorologist can afford to neglect it, and it should find a place on the shelves of all university libraries. It may well do much to increase the interest of physicists generally in the subject and, if so, will prove of inestimable value, for meteorology with its many problems needs greatly to recruit the interest of first-rate minds. But because it may largely spur the minds of readers ready to be enlightened and stimulated, it is the more regrettable that the book gives an impression of hasty compilation, with too dogmatic statement in many places, and on p. 133 it is implied that radiation exchange processes in the atmosphere may on occasion violate the second law of thermodynamics. A second edition will, however, surely be called for, when it is hoped the author will remove such blemishes as now appear. At the same time the publishers, who have made a generally handsome volume, should find a more adequate method of representing surface weather maps ; those given leave practically all the weather data illegible. P. A. S.

Edmond Halley as Physical Geographer, and the Story of his Charts, by S. Chapman. Pp. 15, with 6 figures. (London: Royal Astronomical Society, 1941.) 2*s*. 6*d*. net.

This little monograph recalls to us that Halley is entitled to fame not only as an astronomer but also as the originator of that graphical charting of geophysical data which is now a commonplace in meteorological and geomagnetic studies. Prof. Chapman's title is modest in its claims: not only does he trace the history of Halley as explorer and cartographer, but he also gives us excellent reproductions of three of Halley's meteorological and magnetic charts. A modern map has a perpetual interest for anybody with imagination ; how much more fascinating are these charts, dating back to the times when

> " Geographers on Afric maps
> With savage pictures filled their gaps ",

when the S. Atlantic could be described as the Aethiopick Ocean and the Amazon boldly linked with the Plate ! Not that the interest of the charts is entirely historical : Prof. Chapman shows how modern rectification of Halley's longitudes and incorporation of contemporary data not known to Halley have given us " reasonably reliable knowledge of the distribution of the magnetic declination over a large part of the earth in A.D. 1700 ".

A. H.

Electro-Magnetic Theory, by J. A. STRATTON. Pp. xiv+615. (London: McGraw-Hill Publishing Co., Ltd., 1941.) 42*s*.

In recent years the electromagnetic theory has been applied in fields of research hitherto comparatively little known or, at any rate, studied only for their academic interest. The results of research have led to wide applications of wireless waves in directions beyond their use in broadcasting, to subjects of engineering importance such as antenna design, transmission lines and to cavity oscillations. These new problems call for a change in the study of the principles of electromagnetism and for stress upon branches of it which have hitherto received but scant treatment. They have, in fact, been omitted from textbooks until the last year or two.

These requirements have been met to some extent by a series of publications in the International Series in Physics, and an outstanding contribution is presented in this volume by Professor Stratton.

This work assumes a sound knowledge of the general principles of electricity and magnetism, including the general properties of circuits and some knowledge of the practical aspects of thermionics. It may be said to anticipate that the reader has followed a course of study such as terminates in this country with the introduction to Maxwell's equations. The book begins in the first few lines with the postulation of these equations, so that the reader should know something of Faraday's laws of electromagnetism and their quantitative expression.

The avowed object of the presentation of the subject is to lay less stress upon the study of the stationary state which characterizes Maxwell's treatise and many later works which followed Maxwell's example; a more complete account of variable fields is thus included. The m.k.s. system of units is adopted; i.e. the meter-kilogram-second system with the coulomb as the unit of charge, and the system is explained together with its relation to other systems of units. These units appear to have been adopted more widely in America than in this country, where most workers are accustomed to the mixed system. It would, however, appear that one of the minor changes to which we must accustom ourselves is the spread of this system, which the author points out was first suggested by Maxwell and is not the work of subversive engineers.

In spite of an extended treatment of modern applications, the fundamentals of the theory are fully and carefully presented. Short cuts are avoided and the more subtle and difficult points are adequately dealt with. A little less than half the book is devoted to fundamental aspects of the theory and includes chapters on the field equations, upon field energy and stresses, and upon the electrostatic and magnetostatic fields.

Three extensive chapters on electromagnetic waves follow, and these are devoted to plane, cylindrical, and spherical waves. They include an account of Bessel and spherical harmonic functions such as is required for the purpose of the book. These are treated with a completeness which distinguishes the presentation from a number of works on the subject, and in this respect the author has provided a very useful work of reference.

The book is essentially a theoretical work, but the last two chapters are devoted to practical problems of the highest importance in modern application of electromagnetic theory. These include the well-known derivation of the laws of reflection and refraction at dielectric boundaries and of Fresnel's intensity relations, together with the theory of total reflection and of metallic reflection. Subjects which have been hitherto to some extent or almost entirely omitted from textbooks, such as the theory of the skin effect, waves in hollow pipes, oscillations in a spherical cavity, and the effect of the earth on the propagation of radio waves receive, if not full, at least stimulating attention.

The work is one that will be found continually useful to those who require a sound knowledge of the large-scale phenomena of electricity and magnetism, and the book will be found useful both as an advanced textbook and as a work of reference. H. T. F.

Tables of Physical and Chemical Constants, by G. W. C. KAYE and T. H. LABY. Ninth Edition. Pp. 181. (London: Longmans, Green and Co., 1941.) 18*s*. net.

Apart from an expression of regret at the untimely death of the senior collaborator, it is difficult to find anything new to say about a work so long and firmly established as "Kaye and Laby". So far as present conditions permit, the ninth edition has been brought up to date, but the general character of the book is still that of its 1911 prototype. It is a handy and unpretentious compilation of all the less esoteric physical and chemical constants likely to be required by the average non-specialising physicist.

It is all too easy, in reviewing a work of this kind, to point out deficiencies in selected sections—for instance, x-ray spectral data are quoted only in wave-lengths, and not in wave-numbers, and no indication is given of the transitions to which the emission lines correspond. Such minor deficiencies are, however, mainly if not entirely due to shortage of space, and it would be a much harder task to get a more generally useful set of tables into the same bulk than it is to emit criticisms like this. H. R. R.

A Text-Book of Electricity and Magnetism, by G. R. NOAKES, M.A. Pp. x + 513. (London: Macmillan and Co., Ltd., 1941.) 8*s*. 6*d*.

Although the scope of this book has been mainly decided by the needs of candidates for Higher School Certificate and University Scholarship examinations, it includes much that is outside such syllabuses, this extra material being distinguished as a rule by the use of more closely set type than the rest. Careful attention is given to sources of frequent difficulties and errors and to the graphical representation of experimental results. Examination questions are appended to the chapters, and the text is abundantly illustrated with 311 line diagrams and photographs. The calculus is used throughout.

One of the most striking and most welcome features of the book is its modern outlook ; up-to-date information is given in the descriptions of the more recent developments, and current ideas are freely used in explanations of fundamental principles. The electron-volt, for example, is introduced as early as p. 8, and the Van de Graaff generator at the end of the first chapter.

Electrostatics is treated in the first two chapters, and magnetism in the next two and in three later chapters (XIV, XXI and XXII dealing with magnetic materials, ferromagnetism and geomagnetism) ; current electricity occupies ten chapters (V–XIII and XV), which are followed by one on units and dimensions, one on conduction in gases and five (XVIII–XXII) on more recent developments. Amongst the topics outlined in these later chapters are: determination of electronic charge, mass spectrographs and isotopes, electron tubes (thermionic valve, cathode-ray oscillograph, photoelectric cell, electron microscope, x-ray tube), quantum theory, atomic structure and spectra, radioactivity, nuclear structure and reactions, high-voltage generators, and cosmic radiation.

The Higher School Examinations are as much a test of the schools as of the individual candidates; and there need be little wonder at the high standards attained by candidates from some of the schools if the teaching is on the plane of this first-rate book by the senior science master at Giggleswick and his companion work on Light, which has already been reviewed in these *Proceedings*. It is, in fact, one of the best, if not the best, of the electrical text-books not only for schools, but also for university students in their first and second years. As an introduction to their more advanced studies it can be unreservedly recommended; it is excellent value for a very modest price. W. J.

Wave Motion and Sound, by F. C. CHAMPION. Pp. 67. (London: Blackie and Son, Ltd., 1941.) 5*s*. net.

This book is Part IV of the publisher's *University Physics* series for first- and second-year courses at a university.

Wave motion and the various aspects of sound, such as its generation, intensity, velocity and frequency, are dealt with in the relatively small space of fifty-odd pages. In consequence, the subject matter suffers in places from inadequacy of treatment, and one feels that, to remedy this defect partially, the space allocated to the examples at the end of each chapter could have been used to better purpose. Line diagrams are employed throughout, and on the whole these are well executed, but the clarity of the text is marred in places by errors in punctuation and composition: on page 23 the end correction for the resonance length of a cylindrical tube is wrongly quoted. In the reference to supersonic generators in Chapter IV it is surprising to find no mention of the magneto-strictive oscillator.

The answers and hints for solution of the examples at the end of the book should prove particularly helpful to students. R. W. B. S.

RECENT REPORTS AND CATALOGUES

Mechanical Behaviour of Bitumen, by W. LETHERSICH. (Technical A/T 83, 1941.) Pp. 28. BRITISH ELECTRICAL AND ALLIED INDUSTRIES RESEARCH ASSOCIATION, 15 Savoy Street, London, W.C. 2. 15*s*.

Researches on the Structure of Alloys, by W. HUME-ROTHERY, M.A., D.Sc., F.R.S. Research Report, R.R.A. 562, June 1941.) Pp. 12. BRITISH NON-FERROUS METALS RESEARCH ASSOCIATION, Euston Street, London, N.W. 1. 2*s*. 6*d*.

Electrochemical Analysis Apparatus. (Leaflet GT 1325, August 1941.) GRIFFIN & TATLOCK, Ltd., Kemble Street, London, W.C. 2.

Books in New Condition at Reduced Prices. (Catalogue 595, November 1941.) Pp. 40. W. HEFFER & SONS, Ltd., Cambridge.

Safe Handling of Radioactive Luminous Compound. (National Bureau of Standards Handbook H 27, May 1941.) Pp. 15. U.S. DEPARTMENT OF COMMERCE, Washington.

Testing of Timepieces, by R. E. GOULD. (National Bureau of Standards Circular C 432, August 1941, superseding C 392.) Pp. 27. U.S. DEPARTMENT OF COMMERCE, Washington. 15 cents.

Safety Rules for the Installation and Maintenance of Electric Supply and Communication Lines. (National Bureau of Standards Handbook H 32, September 1941, superseding H 10.) Pp. 177. U.S. DEPARTMENT OF COMMERCE, Washington. 65 cents.

NEW DEL
4 JUL 194
BRARY

THE PROCEEDINGS OF
THE PHYSICAL SOCIETY

VOL. 54, PART 2 1 March 1942 No. 302

THE NATURE AND MEASUREMENT OF WHITENESS *

By J. G. HOLMES, A.R.C.S., B.Sc., F.Inst.P.,
Chance Brothers and Co., Ltd.

MS. received 6 *November* 1941

ABSTRACT. The uniqueness of white is discussed from both the visual and colorimetric points of view. The sensitivity of the eye to departures from white are considered, together with the extent to which the visual mechanism compensates for changes in illumination. A short account is given of the instrumental methods of colorimetry most suitable for the measurement of whiteness.

§ 1. INTRODUCTION AND DEFINITIONS

EVERYONE is familiar with the idea of whiteness relating either to a white substance or to a white light. The perfectly white substance is unique, and is thus different from a yellow or blue substance, which may be found in a wide variety of colours depending on the criterion of observation. Perfectly white light cannot be defined objectively, and it has been found necessary to attempt, and very difficult to obtain, agreement on a standard "white" illuminant.

A *perfect white* substance can be defined as one which reflects all light falling on it. It is a corollary that it should possess no absorption of light at any part of the spectrum. A white substance is one which possesses negligible selective absorption of light at different wavelengths and which, therefore, will not make any appreciable change in the colour of light illuminating it, although it may reflect a smaller amount of light than it receives. *Whiteness*, or the state of being white, relates to the colour or selective absorption of a substance rather than to its reflection factor or overall absorption. It may be pointed out that no reflecting substance possesses colour of its own, its colouring properties depending entirely on the selective absorption which it exerts on the light falling on it. The colour name given to such a substance is usually the complement of the colour which is most strongly absorbed.

The reflection factor of a white substance, that is the ratio of reflected to incident light flux, may vary very widely, from 90% for opal glass to 10% for some textiles, and those concerned with each specific material have to agree as to the border-line between white and grey. The whiteness is, of course, independent of the intensity of illumination under which the substance is observed.

---

* A paper read to the Colour Group on 25 September 1941.

There is no satisfactory theory to account for selective absorption, but the mechanism of "white" reflection is comparatively simple. A beam of light falling on snow crystals is partly reflected at the crystal surfaces and partly refracted through the crystals to suffer reflection at other surfaces, so that eventually it is nearly all diffusely reflected or scattered, undergoing very little absorption and no transformation. The same is true of a suspension, such as milk, opal glass, fog or paint, in which transparent particles are suspended in a transparent medium of different refractive index and the light is reflected specularly at a multiplicity of interfaces. If the two phases had the same refractive index, the materials would be comparatively transparent, like ice suspended in water. Paper and fabrics are white because they consist of non-absorbing transparent fibres in air, powders are white because they consist of minute crystalline or vitreous particles in air; and it is common experience that their reflection factor is controlled by particle size and by the difference in refractive index, being high for dense crystals such as white lead and titania and decreasing when such crystals are immersed in water or liquids of higher refractive index.

A *white light* has been variously defined as "north sky daylight", "mean noon sunlight", or "equi-energy white", and for the purpose of this paper it is sufficient to say that there is a range of white lights from which each may choose the one most suited to his immediate needs, such as the Planckian distribution at 2360° K. for signal lights (BSS. 623—1940) or the C.I.E. illuminant C for artificial daylight (BSS. 950—1941). The colour of a white light is as real and as important as that of a yellow or blue light (Holmes, 1941) and it may be obtained in as many ways, but the requirement for a standard white illuminant is that the distribution of energy through the spectrum should be smooth and constant. This led to the adoption of equi-energy white (C.I.E. 1931) for theoretical purposes and an incandescent source of light with or without a colour filter for both theoretical and practical purposes. The whiteness of such a source is not dependent on its intensity or brightness, although the visual impression may change.

§ 2. VISUAL EFFECTS

The eye is surprisingly critical in respect of white substances, and small departures from white are described as cream or ivory, whereas a wide range of colours would receive the simple colour names red, green, blue, etc. It seems, almost, that there is an ingrained sense of whiteness which has been sharpened by the frequency with which near-white substances are commonly met and can be compared with surrounding objects, so that the uniqueness of white enables it to be recognized, whereas a particular yellow or blue or even a particular grey calls for an effort of memory. This discrimination does not apply to lights which are commonly called "white", and there is just as wide a variation of these as of any other colour, probably because the eye becomes accustomed to changes between sunlight and skylight, electric light, gas light, etc. A white substance can be picked out quite critically in daylight or artificial light, although the actual colour is widely different under the two conditions, and a white substance or a white light will still appear white if the brightness is reduced to about 0·1 foot-lambert where the Purkinje effect is operative but where colour recognition is still fairly critical.

§ 3. COLORIMETRY

The colorimetry of whites is not fundamentally different from that of other colours, but in several ways it is more exacting. Visual matches may be made by additive or subtractive mixing, the former being more flexible but usually less accurate because it will not reproduce the spectral energy distribution of near-white as closely as the latter and is therefore restricted to a much smaller field size than the latter. Departures from white are usually expressed graphically on the C.I.E. chart or R.U.C.S. chart or on some diagram relating specifically to the method of measurement, but in many industrial applications there are three variables (hue, saturation, and reflection factor) which require to be indicated.

The effect of employing different white illuminants may be considerable when examining coloured substances (Nickerson, 1940), but it is less important when examining near-white substances which have no strong absorption bands. It is not generally possible to employ a coloured illuminant to exaggerate small departures from white in the way that one would, for example, employ a blue-green or a monochromatic yellow to exaggerate differences between similar red substances. The quantitative effect of change of illuminant on a white substance or a match between two near-white substances can be calculated exactly, the calculation requiring less precise data than when strongly coloured substances are considered. For most purposes the best illuminant for examining white substances will be the equi-energy source (C.I.E. 1931), which may be obtained fairly closely from a 2848° K. light source in the same way as the C.I.E. illuminants B and C by 1 cm. thicknesses of each of the following Davies-Gibson filters (Davies, 1934):—

Solution E_1

Copper sulphate ($CuSO_4\ 5H_2O$)	2·954 g.
Mannite ($C_6H_8(OH)_6$)	2·954 g.
Pyridine (C_5H_5N)	30·0 ml.
Water (distilled) to make	1000·0 ml.

Solution E_2

Cobalt ammonium sulphate ($CoSO_4(NH_4)_2\ SO_4\ 6H_2O$)	28·440 g.
Copper sulphate ($CuSO_4\ 5H_2O$)	17·840 g.
Sulphuric acid (s.g. 1·835)	10·0 ml.
Water (distilled) to make	1000·0 ml.

The total transmission factor of the combined filter is 19% and the nearest Planckian colour temperature is 5420° K.

Visual tests should always be made with a specimen of ample size, preferably subtending more than 10° at the eye, and with ample illumination, preferably over 10 f.c. If the surface of the sample has a definite texture, the standard method of illumination at 45° and observation at 90° should be adopted.

Visual measurement of the colour of a white light follows the same lines as for a white substance, as it is general practice to illuminate a standard white surface by the light under consideration. Freshly smoked magnesium oxide provides a suitable standard surface, although care is needed in its preparation (Gibson, 1938). Photoelectric measurement does not usually require the intermediary of a standard white surface, as the sensitive element is illuminated directly by the white light or by the light reflected from the surface under examination.

§ 4. INSTRUMENTAL METHODS FOR WHITE SUBSTANCES

The straightforward method of comparing two materials side by side is surprisingly critical, and quantitative methods may be hard taxed to put a figure on differences which are quite obvious. This comparison is, however, generally too simple, and full allowance is not made for such uncontrolled variables as the intensity or the colour of the illumination, differences in texture or differences in background. One of the best means of comparison is the Lummer-Brodhun photometer head, in which the contrast field provides a very sensitive indication of colour difference when any brightness difference has been compensated. Simplified apparatus can be constructed with colourless glass mirrors or a large biprism, by which two similarly illuminated samples can be brought into optical juxtaposition. In all these comparative methods the effect of different illuminants can be closely reproduced by holding suitable colour filters over the eye.

The simplest measurement is that of reflection factor. This may be done visually by comparison with a standard white of known reflection factor, such as magnesium oxide, varying the relative illumination on the two specimens until a match is obtained. For accurate work a photometer bench is desirable, but there are instruments, such as the Hilger photometric comparator, which constitute a more convenient and compact apparatus and provide sufficient accuracy for most purposes. There are many proprietary photoelectric instruments, of which the Morganite reflectometer (Heys-Hallett, 1941) is typical, which give good results, provided the difference between the response curve of the cell and the standard visibility function is not large.

Visual measurement of colour is not easy and may be tackled in three ways:

(*a*) Direct colorimetric measurement.

(*b*) Measurement of reflection factor in several wavelength bands along the spectrum.

(*c*) Spectrophotometric measurement.

In method (*a*) the choice of the colorimeter is important as, for example, a 3-primary additive colorimeter will have such a discontinuous spectral energy distribution that the chromatic aberration of the eye is liable to reduce the accuracy of measurement, particularly with the small field size generally obtainable. Departures from white are usually due to unbalanced selective absorption, which can often be reproduced by subtractive colorimetry as in the Lovibond colorimeter, preferably employing the Schofield system (Schofield, 1939).

Method (*b*) has been considered mostly for assessing the colour-rendering properties of artificial daylight sources, and many papers have been published (Bouma, 1939) on the relative values of different wavelength bands. A British Standard method has recently been described (BSS. 950—1941) and may well be used for estimating departures from white. The apparatus required is the same as for reflection-factor measurement, together with a series of about 8 colour filters isolating different spectral regions. The results may be shown as a graph of reflection factor against the dominant wavelength of each colour filter, and thus much more information is given as to the origin of the departure from " white " than would be given by method (*a*).

Method (*c*) is a more precise form of (*b*), and its main advantage for near-white colours is that it provides a greater number of observations. Visual spectrophotometry is laborious and is generally treated as a last resort in the measurement of whites.

Visual measurements may be simplified by means of a colour amplifier (Pfund, 1920), in which multiple reflection is employed to increase the apparent selective absorption. The amount of departure from white can be magnified up to about ten times, and the gain in accuracy of measurement of a near-white is more than sufficient to compensate the slight doubt regarding the exact power of the magnification.

Photoelectric measurement of colour is still in the early stages of its development, and the three visual methods just described all have their counterparts in photoelectric instruments. Direct-reading colorimetry has been described by Winch (1937), and photoelectric reflection meters can be adapted for the same purpose (Perry, 1938), although these instruments are scarcely sufficiently accurate for near-whites. There are several photoelectric instruments for measuring reflection factor in different wavebands, most of which are based on the reflection meters mentioned above, and precision instruments for this purpose have been described by van Alphen (1939) and by Winch (1940). These instruments are well suited to measurement of whites and give as high an accuracy as the best visual methods. The instrumental readings are best recorded on a diagram of reflection factor plotted against dominant wavelength, but these may be reduced to a simpler description of the sample by calculation of the approximate C.I.E. trichromatic coefficients and of the integrated reflection factor. Photoelectric spectrophotometry has not developed in this country to the same extent as in the United States, where the General Electric recording spectrophotometer (Hardy, 1938, and Wright, 1940) has provided the best method of industrial colour measurement up to the present.

Photoelectric indicating spectrophotometers have been constructed in several laboratories in this country, but their advantages over visual instruments have not been sufficiently great for them to be generally adopted. It is to be hoped that automatic recording spectrophotometry will become better known and used in this country, as this method is probably even better suited to the measurement of whites than of other coloured substances.

REFERENCES

van Alphen, P. M., 1939. *Philips Tech. Rev.* **4**, 66.
Bouma, P. J., 1937. *Philips Tech. Rev.* **2**, 1.
BSS. 623—1940. British Standard Specification No. 623.
BSS. 950—1941. British Standard Specification No. 950.
Commission International d'Eclairage, 1931. *C.R. Acad. Sci., Paris*, 19.
Davies R. and Gibson, K. S., 1934. *J. Res. Bur. Std. Wash.* **12**, 263.
Gibson, K. S. and Keegan, H. J., 1938. *J. Opt. Soc. Amer.* **28**, 372.
Hardy, A. C., 1938. *J. Opt. Soc. Amer.* **28**, 360.
Heys Hallett, C. G., 1941. *J. Brit. Kine. Soc.*, July 1941.
Holmes, J. G., 1941. *Trans. Illum. Engng Soc., Lond.*, **6**, 71.
Nickerson, D., 1941. *Trans. Illum. Engng Soc., N.Y.*, **36**, 373.
Perry, J. W., 1938. *J. Sci. Instrum.* **15**, 270.
Pfund, A. H., 1920. *J. Franklin Inst.* p. 189.
Schofield, R. K., 1939. *J. Sci. Instrum.* **16**, 74.
Winch, G. T. and Machin, C. F., 1940. *Trans. Illum. Engng Soc., Lond.*, **5**, 93.
Winch, G. T. and Palmer, E. H., 1937. *Trans. Illum. Engng Soc., Lond.*, **2**, 137.
Wright, W. D., 1940. *Rep. Progr. Phys.* **7**, 36.

(For Discussion see p. 104.)

THE MEASUREMENT OF NEAR-WHITES IN THE PAPER INDUSTRY *

By V. G. W. HARRISON, Ph.D., F.Inst.P.,
Printing and Allied Trades Research Association

MS. received 6 *November* 1941

§ 1. INTRODUCTION

One of the most important properties of a sheet of paper from both maker's and user's point of view is its colour. Papers of all colours are in use, but of these by far the most important group is the near-whites, on account of their extensive use in printing.

There will be no need to emphasize the necessity for controlling, and hence measuring, the colours of papers. Customers' samples must be matched; and if a papermaker supplies paper according to contract over a period of time, it is essential that the colour of the product should be maintained within fine limits. It looks bad if papers from different batches, nominally the same but differing slightly in colour, are bound up in a book together.

---

* A paper read to the Colour Group on 25 September 1941.

It would be difficult and unprofitable to attempt to set up a rigid definition of a " near-white "; most, however, of the near-white papers in common use are yellows of hue (dominant wavelength) 5500–6000 A., about 3–15 % saturation (excitation purity), and 60–95 % brightness (visual efficiency). A typical curve for a pre-war newsprint is shown in figure 1. The paper is made largely from mechanical wood and contains little or no added dye or pigment. Figure 2 shows a spectro-photometric curve for a paper which has had a small quantity of a blue-green dye added to the stock in order partly to neutralize its yellowish colour.

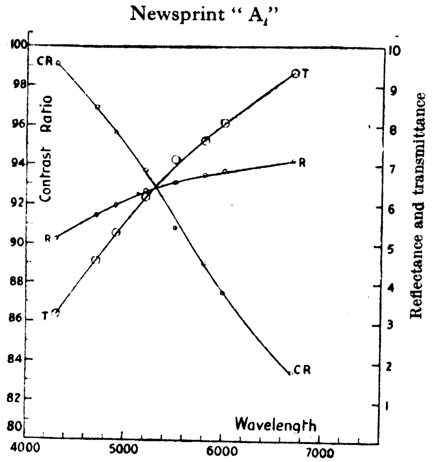


Figure 1. Curves for a typical undyed newsprint. 4000 to 7000 A.
RR = reflectance;
TT = transmittance (arbitrary units);
CR.CR = contrast ratio (a measure of opacity);
all plotted against wavelength of the incident light.

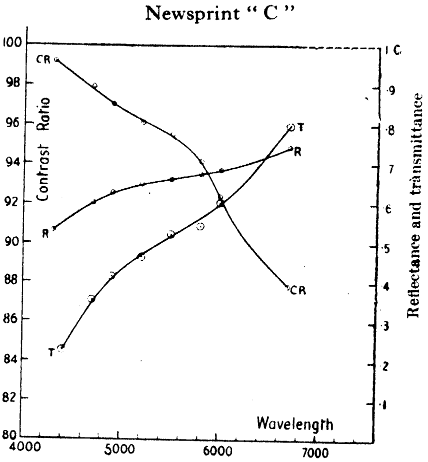


Figure 2. Curves for a newsprint containing a small amount of blue-green dye. 4000 to 7000 A.
RR = reflectance;
TT = transmittance;
CR.CR = contrast ratio;
all plotted against wavelength.

True neutral whites or greys are rare in paper. The papermaker accordingly adopts subconsciously a *natural paper white* (that is, the colour of untreated pulp or paper) as a working standard: papers of lower saturation than this he calls *blue-whites*, and papers of greater saturation he calls *cream-whites*; though all are, in reality, desaturated yellows. The standard of *natural white paper* almost certainly varies from one maker to another, and even with the same maker it probably varies from day to day. In any case, he would have difficulty in defining it or providing a typical example (Judd, 1935 and 1936). The terms blue-white and cream-white are thus more relative than absolute in the paper trade.

The papermaker and paper-user do not, in general, fully realize the threefold nature of colour (even for near-whites), and often insist on a single measure of brightness or whiteness (usually as a percentage) to define the colour of their products. This has led to confusion, and the paper trade's conception of brightness and whiteness needs careful examination, since it differs in important respects from the usage in other trades and, in particular, from the scientific usage of the term *brightness*.

§ 2. BRIGHTNESS AND WHITENESS

The term *brightness* has many possible meanings. In the paper trade it seems to be used to denote the relative quantity of light (independently of its colour) emitted from one surface as compared with another, estimated visually *without the aid of an instrument*. This usage (which for want of a better term I shall call visual brightness) is vague and hardly allows of quantitative specification. One can, however, usually grade papers in order of visual brightness and give a rough indication of the relative magnitudes of the visual-brightness differences. The order of papers graded in this way does not, in general, agree with their order as determined by reflectometer measurements.

Whiteness is one of a number of terms, usually terminating in -ness or -ability, to which it is difficult to give a satisfactory definition, although a certain vague meaning is easily apparent. *Whiteness* seems to be used in at least three distinct senses.

A. To some, *whiteness* is almost, if not entirely, synonymous with " visual brightness ", the " whiter " paper being the one which *appears* to reflect more light.

B. The term *whiteness* may also be used to denote the extent to which the colour of a paper approaches neutrality, regardless of the amount of light reflected from the surface. In other words, whiteness may be regarded as an inverse measure of saturation.

C. A perfect white surface is one which reflects *all* the light which falls upon it without selective absorption. *Whiteness* may therefore be taken to mean the extent to which a colour approaches that of perfect white, and such a definition would take into account not only the colorimetric brightness of the colour under consideration, but also the extent of its departure from neutrality. In this connexion it is not easy to give any quantitative measure of the extent of departure of a colour from perfect white. However, colours are defined by three independent variables and hence may be represented by points in a three-dimensional diagram. The points corresponding to all real colours fall within, or on the surface of, a solid, the apex of which represents perfect white. Colours that most nearly approach white are represented by points nearest the apex of the solid figure; while black is represented by the base of the solid. The most logical measure of the whiteness of a colour would involve the length of the line joining the point representing the colour to the apex of the figure; this length could be found by calculation, but it would depend upon the system of co-ordinates adopted for plotting the figure. In any case, no single number will define a colour (if we exclude quite arbitrary systems), and it will always be possible to find a large number of distinguishable colours all of which have the same degree of whiteness, however this term be defined. For these reasons I suggest that it would be better to abandon the use of the word whiteness altogether in scientific and technical publications.

§ 3. "VISUAL BRIGHTNESS"

To return now to the subject of visual brightness, the published trade literature thereon is mostly unsatisfactory, a fact due probably to the vagueness of meaning of the term and to the mistaken idea that a single figure will suffice for grading a paper. Many " brightness meters " have been described; most of them are means for determining the reflectance of a sample in white or coloured light under more or less arbitrary conditions, and it is usually agreed that a reflectance measurement made in *blue* light gives results in closest agreement with visual judgment (Davis, 1935; Hunter, 1937; Farebrother, 1937).

On this basis depends the General Electric Reflection Meter (Davis, 1935), which is an instrument for determining the reflectance of a specimen when illuminated at 45° and viewed normally in light of wavelength 4500 A. This reflectance is often known in America as the brightness of the paper. Very careful calibration of the instrument is necessary, as the results depend upon the spectral sensitivity of the photocell used, and several other factors. The instrument has been found useful for comparing closely similar batches of paper, but otherwise it is of limited value. M. N. Davies (1935) points out that this use of the term " brightness " " appears to be peculiar to the paper industry. It is probably closely approximated by, if not synonymous with, the term whiteness used with reference to white textiles. It appears to be much more closely associated with the quality purity than with the quality brightness as the latter two terms are used in spectrophotometric colorimetry."

This apparently general preference for papers of lower saturation, even if this entails some loss of colorimetric brightness, is remarkable. At first I thought it was probably due to unconscious bias on the part of trained paper graders, who may prefer highly-dyed, nearly neutral papers to yellower papers as being of better " quality ", and therefore, perhaps, by association, of greater brightness. I therefore carried out some simple experiments on 20 members of the staff of Patra (the Printing and Allied Trades Research Association). Most of these observers had had no previous experience in grading papers and were unacquainted with the literature of the subject. The question of bias may thus be eliminated.

The observers were asked to choose the brighter from pairs of sheets placed in turn before them. Eight sheets were used, giving a total of twenty-eight pairs, all of which were considered independently. The sheets were placed on a table in front of the observer and illuminated by diffuse light coming from a window on the left. No other observers were present during each experiment and no comparison of results was made until all opinions had been recorded.

A full description of the results has already been given in Patra Research Report No. 2, and I shall give only a summary of the main points. Figure 3 shows the approximate spectro-photometric curves of two papers marked for reference " 8991 " and " Bond ". It will be seen that 8991 has the lower reflectance throughout the visible spectrum, but the curve of Bond has the steeper slope. Colorimetrically, Bond is the brighter paper; it also has the greater saturation—that is, it is yellower in colour than 8991. Yet of the twenty observers, eighteen expressed a decided preference for 8991 as the brighter, while two were uncertain. No one voted for Bond as the brighter. Results with the other pairs left no doubt that the principal factor influencing judgment was the slope of the spectrophotometric curve (that is, the saturation of the sample, rather than its colorimetric brightness.) In the paper mill it is common practice for the beaterman to add a little blue dye to the stock in order to " brighten " the paper. From the physicist's point of view, this is nonsense; the dye can only help to neutralize the natural yellowish colour of the pulp, and in doing so it must inevitably reduce the reflection factor of the finished paper: however, these experiments show that, as far as visual judgment goes, the beaterman is behaving quite correctly.

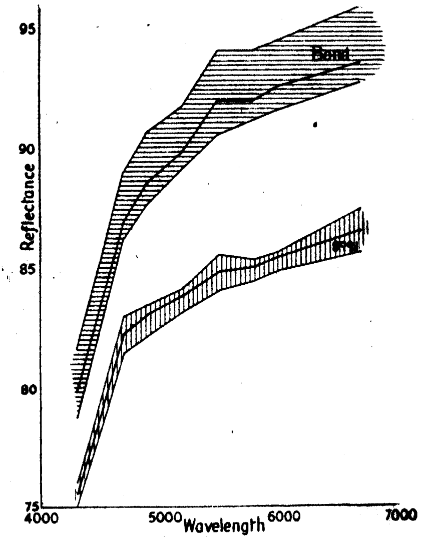


Figure 3. Approximate spectrophotometric curves of two papers. 4000 to 7000 A. The widths of the shaded bands indicate the variations in reflectance observed over the surface of one and the same sheet: the continuous lines near the middle of each band represent the mean values. 18 observers out of 20 considered 8991 to be the *brighter*.

From a study of the results, it was possible to arrange the eight test papers in order of their " visual brightness ". Reflectance measurements were also made on them for light of different wavelengths. Table 1 shows a summary of the results. In the first column are the papers arranged in order of " visual brightness ", starting with the brightest at the top; in the second, third, and fourth columns the papers are arranged according to reflectance measurements made at wavelengths of 6700, 5500 and 4300 A. respectively. It will be seen that reflectance measurements made in red and green light provide a very poor measure of the visual grading of samples. In blue light the agreement is better, though it is still far from good: this gives some experimental justification for the use of the G. E. Reflection Meter already mentioned.

Table 1

Order of sheets according to			
Visual brightness	Reflectance at		
	6700 A.	5500 A.	4300 A.
1	Bond	Bond 8992	8974
8974	8992	8974	1
9138	8974 1	1	8992
8292	8292	8292	8292
8992 9089	9138	9138	Bond
8991	9089	9089	9138
Bond	8991	8991	9089
—	—	—	8991

In table 2, I have compared the orders of the sheets according to " visual brightness " and also according to $2B-R$ and $R-B$, where R and B are the reflectances in red and in blue light respectively. It will be seen that the agreement is now much better, particularly with the $R-B$ formula. (The $2B-R$ formula has been given by Samuelsen, 1937.) It is clear, therefore, that what I have

Table 2

Order of sheets according to		
Visual brightness	$2B-R$	$B-R$
1	8974	8974
8974	1	1
9138	8992	8292
8292	8292	8992 9089
8992 9089	9089	9138
8991	9138	8991
Bond	Bond	Bond
—	8991	—

called *visual brightness* is closely related to the *saturation* of the colour, and *not* to the quantity of light reflected; indeed, within limits, the colorimetric brightness seems to be of comparatively small importance in affecting estimates of visual brightness. This result, of course, is valid only for near-white colours.

It thus appears, from these experiments, that the eye is, for near-whites, more sensitive to small changes in saturation than to small changes in colorimetric brightness. The effect of a small increase in saturation is, to most observers, to produce a decrease in visual brightness. As the saturation is increased, the colour obviously becomes yellower to all observers: the interesting fact, however, is that the initial increase in saturation, if not too great, makes itself apparent rather as a decrease in *brightness* (visual) than as an increased yellowness. This may be illustrated by the sheet marked Bond, which almost all observers agreed was dull in comparison with the others. Several, however, failed to notice any particular yellow or cream tint about it, but described it as comparatively grey.

It is apparent from the foregoing that single brightness, whiteness or reflectometer measurements are only of very limited value in specifying the colours of near-white papers. Near-whites are still colours, and as such they can only be adequately defined in terms of the C.I.E. system. The accuracy required and the type of instruments most suited for the purpose must therefore next be considered.

§ 4. METHODS OF MEASURING OR DEFINING THE COLOUR OF PAPER

A. The making of colour matches with samples in a colour catalogue. The obvious objections to such a system of colour notation are: (1) the selection of the standard colours for the catalogue must be arbitrary; (2) it is difficult to reproduce the standard colours accurately; (3) it is left to the individual judgment of the observer as to which of the standard colours is the nearest match to the sample, and opinions here can vary to a surprising degree; and (4) it is difficult to ensure permanency of the standard colours, and their fading, when it occurs, cannot easily be detected.

B. Matching the colour of the specimen with that of a combination of filters chosen from a set of various " depths " of three primary colours. The Lovibond tintometer is of this type.

C. The visual trichromatic colorimeter, such as the Guild and Donaldson colorimeters.

D. The photoelectric trichromatic colorimeter. Several of these have been designed and constructed. They vary from simple reflectometers of doubtful value, making use of arbitrary red, green and blue filters in front of the photocell, to elaborate precision instruments, of which the Blancometer is probably the best-known example.

E. The monochromatic colorimeter, in which the colour of the sample is matched visually with a mixture of white light and one of the colours of the spectrum. This instrument is used mainly for research on colour vision, and I am not aware of any cases in which it is in use as an industrial colorimeter.

F. The "abridged" spectrophotometer, or Reflectometer, which makes use of a set of filters for changing the wavelength of the incident light.

G. True spectrophotometers, in which the sample is illuminated by monochromatic light.

I do not propose to give further consideration to methods A and E. Whether the other methods can be considered suitable or not depends mainly on what degree of tolerance in colour measurements is permissible in commercial work.

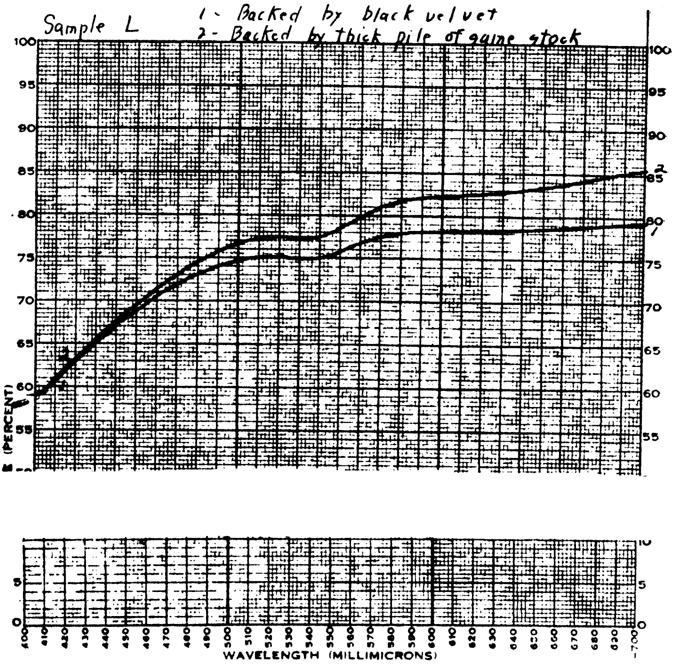


Figure 4.

There is little published information on what degree of accuracy papermakers require. It is, however, probably related to the least perceptible difference in chromaticity obtainable under the best conditions of observation; and in this connexion it must be remembered that the sensitivity of the eye when one is observing large sheets side by side in a good light is not necessarily the same as its sensitivity when one is working with the poor light and narrow 2° field of a colorimeter. Judd (1936) gives a tolerance figure of 0·0005 in the trichromatic units. This is very small.

In order to gain information as to the degree of sensitivity required for commercial work, six batches of paper were obtained by Patra from a papermaker.

These were taken from a single run in the mill and were identical except for small colour variations. These variations were so large as to make it necessary to sort the paper into six distinct batches, so that the colour differences clearly fell outside the commercial limits of tolerance. Samples of these papers and data for them have been circulated in a Patra Research Report (No. 3). By courtesy of the Research Laboratories of Interchemical Corporation, New York, spectrophotometric curves were taken for these papers on the Hardy recording spectrophotometer. A typical curve is shown in figure 4. From these curves the C.I.E. trichromatic units and brightness factors were calculated for Illuminant B by the thirty-selected-ordinate method: the results are summarized in table 3.

Table 3

ILLUMINANT "B". SELF BACKING

Paper	Tristimulus values.			Trichromatic coefficients	Brightness ($\bar{Y}$)
	$\bar{X}$	$\bar{Y}$	$\bar{Z}$		%
C	78·52	78·84	58·00	$0{\cdot}3646X+0{\cdot}3661Y+0{\cdot}2693Z$	78·8
CP	76·66	76·93	56·97	$0{\cdot}3641X+0{\cdot}3654Y+0{\cdot}2706Z$	76·9
L	79·24	79·69	59·99	$0{\cdot}3620X+0{\cdot}3640Y+0{\cdot}2740Z$	79·7
BB	79·63	80·05	60·50	$0{\cdot}3617X+0{\cdot}3636Y+0{\cdot}2747Z$	80·0
BC	78·53	78·77	57·26	$0{\cdot}3660X+0{\cdot}3671Y+0{\cdot}2669Z$	78·8
PR	76·12	76·32	56·03	$0{\cdot}3651X+0{\cdot}3661Y+0{\cdot}2688Z$	76·3
D	76·55	76·83	56·89	$0{\cdot}3640X+0{\cdot}3654Y+0{\cdot}2706Z$	76·8

The papers marked CP and D were identical, but were sent for test as different batches in order to test the performance of the instrument. I will not vouch for the absolute accuracy of the figures (this depends upon the calibration of the instrument, which, of course, we had to take on trust), but I do consider that they give a fair measure of the sort of colour differences that we are called upon to detect. It will be seen that for the extreme papers of the range the differences in the coefficients of X and Y are only about 0·004, and the extreme brightness differences are only about 4 %. Thus to distinguish between the papers with certainty involves measurements to a considerably higher degree of precision than this; indeed, from the figures it appears probable that it will be necessary to estimate the fourth figure of the trichromatic units and the brightness factor to 0·2–0·3 %. This is well outside the limits of visual colorimetry.

As a check on these results, we attempted to measure the colours of these papers on a Donaldson colorimeter: we found it quite impossible to separate the papers with any degree of certainty. A member firm also made measurements for us with a Lovibond tintometer, and reached similar conclusions. On the other hand, a laboratory reflectometer fitted with a series of eight "monochromatic" filters showed clear differences between the papers, and also yielded curves in fair agreement with those given by the Hardy instrument (figure 5); unfortunately, however, these curves were partly characteristic of the instrument as well as of the paper, and consequently could not be made to give results on an absolute basis.

From these results we have concluded that visual colorimeters are unsuitable for the measurement of near-whites in the paper industry. Personal errors are much too large and they are insufficiently sensitive. I believe it is a fact that the eye is more sensitive to small differences of chromaticity when the observer is free to compare large samples side by side in a good light than it is when vision is restricted to the weak light and narrow field of a colorimeter. Thus the failure of these instruments is due not to any fault in their design or construction, but to the characteristics of the human eye.

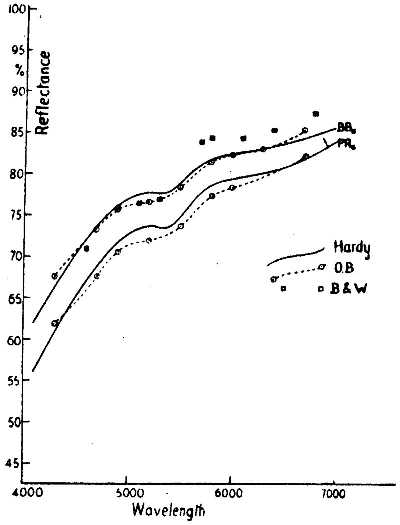


Figure 5. "Hardy"-spectrophotometric curves given by Hardy recording spectrophotometer. 4000 to 7000 A.

"O.B." and "B & W": points given by two "abridged" spectrophotometers.

The choice of instruments is thus restricted to photoelectric colorimeters or spectrophotometers. Each of these classes of instrument has its own particular advantages and disadvantages.

§ 5. REQUIREMENTS OF AN INSTRUMENT FOR INDUSTRIAL PURPOSES

In the ideal case it should be possible to specify the colour of a piece of paper in terms of the C.I.E. system, and to reproduce that colour within commercial tolerance limits in other works without the need for an exchange of samples. If one orders a shaft to be turned to a certain length and diameter one can be certain, provided the work has been properly carried out, that the shaft will fit the job for which it is intended; there is no need for one to send a sample shaft for the engineer to match. A similar state of affairs should be possible with colours. However, this involves measuring the trichromatic coefficients and brightness factors to a very high degree of absolute accuracy, so high, indeed, that special difficulties are introduced. One may therefore decide to abandon hopes of getting the requisite degree of absolute accuracy, and to rest content with producing an instrument giving more or less arbitrary readings for internal use in the works only.

If the second, less ambitious, course is followed, the main requirements are: (1) adequate sensitivity, (2) stability, (3) the results, even if arbitrary, should be clearly related to what it is desired to measure. The second requirement is necessary because if one makes measurements on the same paper over a period of time and variations are noticed, one does not want to have to analyse the results to find out whether the variations are due to changes in the paper, or to the erratic behaviour of the instrument. The third requirement is necessary because it is easy to construct so-called colorimeters which, although they are sensitive and give repeatable readings, are liable to give misleading results; for example, a badly designed colorimeter may give identical readings for two papers which differ perceptibly in colour.

§6. SPECTROPHOTOMETER OR PHOTOELECTRIC COLORIMETER

If one is satisfied with an instrument intended for comparative results for use in the works only, fairly satisfactory approximate spectrophotometers and colorimeters can be constructed or obtained without much difficulty.

The type of " abridged " spectrophotometer in common use is essentially a photoelectric reflectometer equipped with a number (seven or eight is convenient) of narrow-band filters. With such instruments, sensitivity, coupled with a reasonable degree of stability, can easily be obtained, and they are not costly. They are particularly useful if the effect of adding dyes to a paper stock is being studied. On the other hand, the readings cannot be converted into absolute (C.I.E.) units, and they retain an arbitrary character due to instrumental factors that cannot be completely eliminated from the results. These instrumental errors rapidly become greater as the colour to be measured departs from white.

The term *photoelectric colorimeter* covers a wide range of instruments, from crude reflectometers making use of arbitrary red, green and blue filters to highly-developed instruments of the " Blancometer " type. Most of the cheaper type are decidedly unsatisfactory; they are insufficiently sensitive, and the readings, if not quite arbitrary, contain large instrumental factors for which it is difficult to make proper correction. A few of the better types, however, can be very useful for routine control purposes. In them the instrumental errors have been reduced to comparatively small dimensions, and their sensitivity and stability are adequate. They have the advantage that they furnish approximate readings in terms of C.I.E. units *directly*. On the other hand, the better instruments are costly, they give results for one type of illuminant only (unless specially constructed to carry additional filters for each extra illuminant required), and for some work a mere statement of colour in terms of trichromatic units gives much less information than the complete spectrophotometric curve. The C.I.E. units can always be calculated from the curve (though the calculation is tedious), but the converse does not apply.

If, however, we are aiming at absolute accuracy within the limits of commercial tolerance, the problem is much more difficult. Only the best instruments need be taken into consideration, and their cost will inevitably be high. The Blancometer may be capable of giving the necessary accuracy provided it is very carefully and frequently calibrated. The " abridged " spectrophotometer here is useless. Since it is thus necessary to employ an expensive and probably delicate instrument, it is most economical to adopt the instrument which gives the fullest possible information about the colour to be measured. This instrument is undoubtedly the spectrophotometer. A recording spectrophotometer, such as Hardy's,* has obvious advantages, but it is not absolutely essential.

I can see little possibility of the general adoption of recording spectrophotometers by private firms. For one reason, to all but the largest firms the cost of the installation would probably be considered prohibitive; for another, unless the spectrophotometers were carefully and intelligently used, a state of chaos would almost certainly arise and the use of instruments probably abandoned altogether.

There is, I think, a way out of the difficulty. Most papermakers can manage very well provided they have working standards which they have to match. This can be done by eye or with the help of one of the " abridged " spectrophotometers or photoelectric colorimeters already discussed. The trouble is that paper is always liable to discolour with age. Slight discoloration in an alleged standard is difficult to detect, and unless noticed and corrected, a mill may easily find itself producing progressively yellower and yellower papers. Moreover, there are at present no accepted standards to which to work: the pre-war procedure used to be for a customer to send in a sample of paper to the mill and expect the mill to match it exactly, regardless of whether the mill had a close, but not exact, match already in production as a stock line.

Most printing could be done without disadvantage on one or other of a limited number of standard papers. If the colours of these standard papers could be agreed upon, a very useful simplification would result. The next step, in my view, would be for a Paper Testing Centre to be set up, which would undertake to provide spectrophotometric and colorimetric data for paper samples, and to assist in their interpretation in cases of difficulty. The mill would then send its standard lines along regularly for test, say once a month. In this way any alterations in the colours of the samples would be detected before they had time to become serious, and if necessary, new standards could be prepared. Once the mill was satisfied that the standards were satisfactory, the production of papers to match the colour standards would be found straightforward. Papers of other colours would then only be made to special order.

---

* **An account of the design, construction, calibration and operation of the Hardy recording spectrophotometer will be found in a series of papers in the *J. Opt. Soc. Amer.* 1938, 28 (10), 360-386.**

It is, I believe, only along the lines of standardization and co-operative testing that the present difficulties and wastage of materials involved in the matching of customers' samples can be overcome.

REFERENCES

DAVIS, M. N., 1935. *Paper Trade J.* **101** (1), 36.
FAREBROTHER, T. H., 1937. *Proc. Tech. Sect. Paper Makers' Assoc.* **18** (1B), 147.
HUNTER, R. S., 1935. *Paper Trade J.* **100** (26), 37.
JUDD, D. B., 1935. *Paper Trade J.* **100** (21), 40 ; 1936. *Ibid.* **103** (8), 38.
JUDD, D. B., 1936. *J. Opt. Soc. Amer.* **26**, 421.
SAMUELSEN, S., 1937. *Proc. Tech. Sect. Paper Makers' Assoc.* **17** (2), 297.

(For Discussion see p. 104.)

THE WHITENESS OF CINEMA SCREENS*

BY C. G. HEYS HALLETT,
Morgan Crucible Co., Ltd.

MS. received 6 *November* 1941

§ 1. INTRODUCTION

WHEN black and white films are projected, the apparent colour of the screen is of no practical importance; indeed, some producers use either lavender or sepia stock on occasions in order to enhance certain dramatic effects. The advent of colour changed all this, and the whiteness of the screen is now of first-class importance if colour distortion is to be avoided. An immense amount of work has been expended on the development of colour processes such as Technicolor and Dufaycolor, but, although release prints of really high quality are now available, little has been done to provide undistorted projection. A start has, however, been made on a survey of projection conditions in a large number of cinemas.

The special instrument, which was developed for this work, consists of a lens which throws an image of the screen on to a photo-electric cell connected to a galvanometer which thus records the total amount of light received from the screen. Arrangements are made for the insertion of suitable masks to enable part only of the screen to be examined, and also for the insertion of a tri-colour set of filters to enable measurements of colour to be made.

---

* A paper read to the Colour Group on 25 September 1941.

The instrument was primarily designed for use by service engineers when testing and adjusting the projection equipment, and is therefore invariably used from the projection room (see figure 1) with shutter running but no film in the gate.

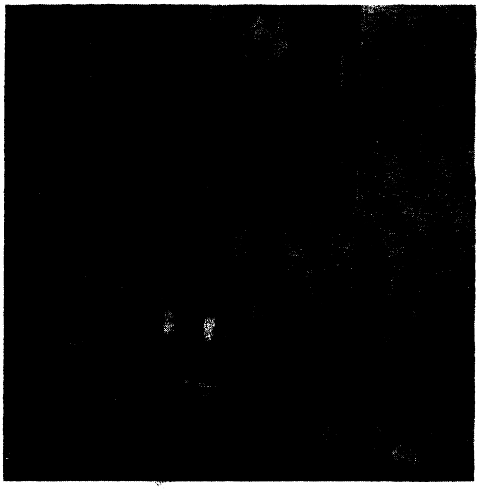


Figure 1. The reflectometer mounted in the projection room.

§ 2. MEASUREMENT OF BRIGHTNESS

As the image is smaller than the active surface of the photo-electric cell, it is necessary to correct the readings by means of a chart to obtain a measure of brightness. Before the development of this instrument, it was customary to measure the incident light on the screen, and most people connected with projection have some idea of the values to be expected, but their knowledge of photometrical units is generally insufficient to enable them to appreciate the difference between foot-candles and foot-lamberts, and endless confusion resulted. To overcome this, the readings taken in a large number of cinemas were examined and, after eliminating unusual and obsolete equipments, an average was obtained. This average was then used as a datum and all readings were expressed as a ratio; thus, if a cinema showed a reading of 66 scale divisions, this was divided by the datum reading of 60 and the cinema was then said to have a *specific brightness* of 1·1, indicating that it was 10% above the datum (which is, incidentally, approximately 5 foot-lamberts).

Table 1

Case	Class of house	Specific brightness	Remarks
31	Large West End	0·52	
16	Large West End	1·21 1·65	Change due to re-treatment of screen
95	Large West End	0·80	
91	Modern large town	0·76	Modern arcs under-loaded
89	Modern large town	0·82	Same box equipment as 91, but fully loaded
74	Small country town	0·78	A.C. high-intensity
75	Small country town	0·28	Low-intensity
93	Suburb	0·16	
86	West End News Theatre	8·60	Same equipment as 89
56	West End News Theatre	4·43	Beaded screen
87	Renter's Theatre	0·314	Low-intensity arcs
96	Renter's Theatre	6·65	Obsolete arcs
36	Studio Rush Room	4·87	
14	Two Rush Rooms in Studio	1·67 2·61	
37	Colour Film Laboratory	2·62	As found
		1·15	After adjustment with Reflectometer
A.12		0·24	Dirty screen, modern lamps
A.16	Kinemas in Sydney and Melbourne	0·98	
A.19		2·0	
A.60		0·59	A.C. high-intensity

The above table shows the magnitude of the variation in brightness which is encountered in practice.

§ 3. COLOUR MEASUREMENT

Two methods of expressing results are in use. Prediction of visual effect can obviously only be based on comparison of the light under review with a standard. We were, therefore, forced to adopt a standard at the commencement of our work, and the standard selected might be described as an average of high-intensity projection arcs.

(a) *Equal-green method*

This method of comparing two colours depends on eliminating one of the filter readings, always the green, by expressing the set of readings as a percentage of the green reading. Even then the set of readings for a kinema would be quite meaningless to everyone, and recourse is therefore had to the solution worked out for screen brightness. That is, all readings are expressed as a function of the datum and termed *specific colour*.

The example given in table 2 shows readings of a kinema which has a light which is 31 % high in red and 25 % deficient in blue. It is possible with a little practice to get some idea of what the visual effect will be, but as two specific colours are involved it is not easy. This example is one of the simpler ones, but sometimes both blue and red specific colours exceed unity. A method of reducing the description to a single expression has therefore been evolved.

Table 2. Specific colour

Colour	Actual readings	Expressed as % of green	Standard colour	Specific colour
Red	6·3	59	45	1·31
Green	10·7	100	100	1·0
Blue	4·0	37·5	50	0·75

(b) *Dominant hue*

Knowing the specific colours, calculate the percentage of the total represented by each, thus:—

$$\begin{array}{ll} P_r & 1{\cdot}31 \\ P_g & 1{\cdot}0 \\ P_b & 0{\cdot}75 \\ \hline & 3{\cdot}06 \end{array}$$

therefore

$$\% R = \frac{1{\cdot}31 \times 100}{3{\cdot}06} = 42{\cdot}8$$

$$\% G = 32{\cdot}7$$

$$\% B = 24{\cdot}5$$

Similar treatment of the standard colour (P_r 1·0, P_g 1·0, P_b 1·0) obviously gives 33⅓ for each.

These colours are then plotted on the wheel diagram (figure 2). Standard colour is represented by the point at the centre of the diagram, since if we plot 33⅓ units towards green, turn 60° and plot 33⅓ units towards blue, followed by 33⅓ units in the direction of red, we return to the point from which we started. The other set is then plotted in a similar manner, and the triangle fails to close, indicating a colour difference of 15·6 % orange, the visual effect of which is easily pictured in the mind.

When using this diagram to compare two non-standard lights we proceed as follows:—

First, multiply each colour reading of one light by the requisite factor to bring it to 33⅓ and thus make it the centre of the diagram.

Now multiply each colour reading of the other light by the same factor as was used for the same colour of the first light. Add up the three results, and multiply each by (100/total). The results thus obtained will add up to 100, and are the percentage readings required for the colour wheel. Note that in this case it is not necessary first to calculate the specific colours.

This method is not very accurate if the lights are far off white, but is sufficiently close for the purposes under review.

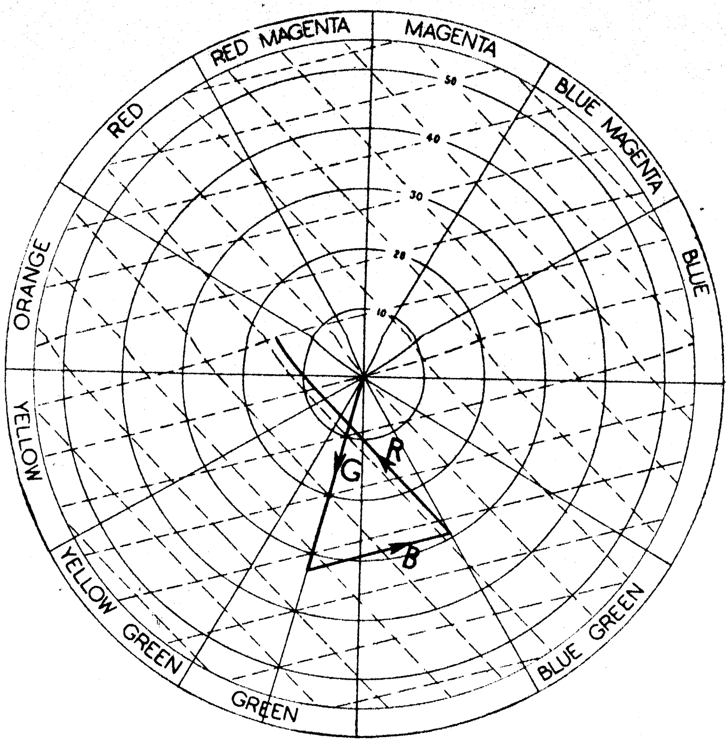


Figure 2. Method of plotting dominant hue on colour circle.
(Reproduced by the courtesy of the British Kinematograph Society.)

It is now possible to study the variation in colour encountered in kinemas, and this can best be done by considering the various links in the chain connecting the arc and the eye separately.

§ 4. VARIATION OF COLOUR

The colour of projection arcs varies for a number of reasons. The shaded area in figure 3 shows these variations for a representative modern high-intensity mirror arc designed to operate round about 50 ampères. The lamp was tested over a range of 10 ampères with several grades of carbon of different makes. The lamp was properly handled throughout the test, and the shaded area therefore represents the envelope which will be obtained when the lamp is run strictly in accordance with the maker's intentions. Cases of under- or over-loading will, however, occur in practice.

The difference between high intensity and low intensity is most marked. The point A in the diagram indicates the average of low-intensity mirror arcs when correctly run at approximately full rating. Reduced loading will, of course, make the light even redder.

The effect of unusual design is shown by point C, which indicates the difference between the Stelmar arc and the representative lamp represented by the shaded area, while B is the point for an A.C. high-intensity projection arc.

Some attempts have been made to use metal mirrors instead of glass, and point F, comparing a rhodium mirror with a glass one, is therefore of interest.

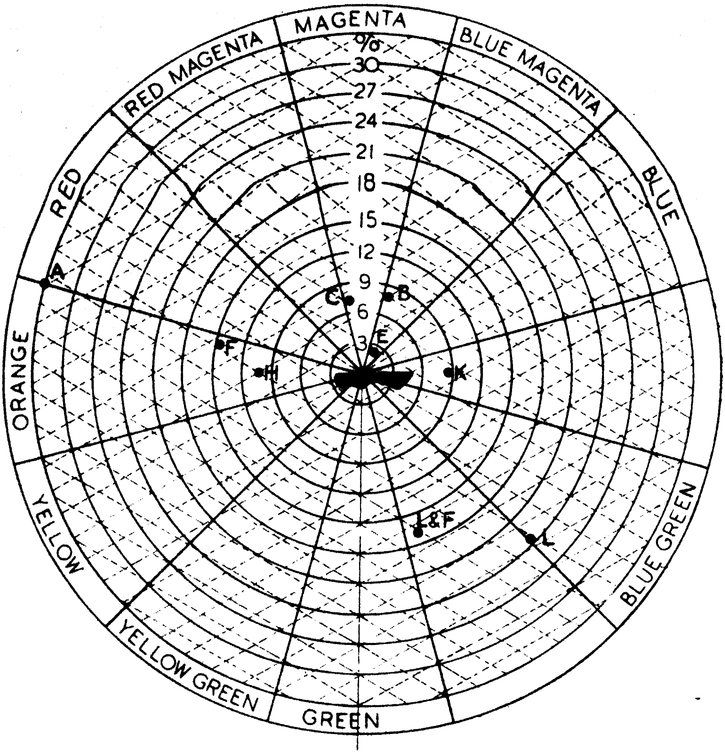


Figure 3. Examples of readings plotted on colour circle.
(Reproduced by the courtesy of the British Kinematograph Society.)

Screens have a powerful influence as illustrated by point H, which shows the change in colour during the life of the screen. Point K is for a beaded screen, while all other types of screen tested lie, when new, between 2·5 % orange and 3·5 % blue-green. Point E shows the change effected by re-treating the screen in a particular kinema.

§ 5. DISTORTION OF THE PROJECTED PICTURE

Screen brightness and the apparent colour of the screen each have their own distorting effect on colour films. Experiments on the effect of brightness have been carried out in the Morganite Laboratories, and it was found that with both black-and-white and colour films, the picture improved progressively as the brightness increased, that a definite maximum was reached with black and white at a specific brightness of about 10, above which graininess became troublesome, but that even at a specific brightness of about 50 no limit could be found for colour films. The effect of low brightness was much more serious on colour than on black-and-white because more colour distortion resulted. At specific brightness below 1·0 all colours appear to be mixed with brown, thus giving the picture an " old master " appearance. Screen deterioration, therefore, may distort the apparent colour in two ways, firstly by loss of reflectivity and secondly by introducing an orange hue.

Attempts have been made to correct the apparent colour of the screen in various ways, but these all seem to be doomed to failure because they all reduce the screen brightness, and as nearly all screens are inadequately lit, the distortion introduced by this reduction usually exceeds the gain in truth due to correction of colour.

The use of blue glass mirrors is an example of the attempts which have been made. The object was to correct the typical orange hue of low-intensity arcs to the blue of the modern high-intensity. A particular mirror of this type was compared with a clear glass mirror and found to reflect 83 % through the blue tricolour filter, 60 % green and 49 % red.

Changing the clear glass mirror of a low-intensity arc for this mirror altered the dominant hue of the light incident on the screen from 21 % red to 8½ % orange, which is still far from perfect. This improvement was obtained at a cost of 39 % of the screen brightness, and as low-intensity arcs have, as their name implies, only about ⅓ of the light output of a high-intensity arc, this cost is prohibitive.

§ 6. CONCLUSION

Unfortunately, insufficient work has been done to enable the requirements of screen brightness and colour to be laid down. Release colour prints are now so good that, when projected under the conditions for which they are made, the colour reproduction is perfect within the discrimination of the average audience.

The deviations of dominant hue usually encountered are within the scope of the colour processes, and it may therefore be said that, as a release print must be shown in any cinema, the prime necessity is standardization, and exactly what is standardized is of secondary importance.

The problem of brightness is, however, not so easily settled, because the great majority of the cinemas in this country have inadequate brightness, and it is therefore difficult to see how the rendition of colour films can be improved without wholesale re-equipment.

DISCUSSION*

Mr. J. W. **Perry**. One is reminded by this sequence of papers of the dictum of Prof. Freundlich that " Science consists in approximations ", which may here be understood in relation to the growth of colorimetry in both a qualitative and a quantitative sense. Recent applications of colorimetry have made progressively increasing demands upon accuracy as the conflicting effects of physical requirements and conditions of accuracy come to be more fully recognized with the growth of knowledge of their individual relationship to the discrimination limen. The earlier stages of this evolution may be traced through application of colour names and samples.

---

* On the three papers above.

Whiteness in a physical sense may be regarded as a quantified colour name, and in the sense that colours may be said to " have whiteness ", this property would be possessed by the colours represented within a certain spatial region of a solid colour chart. Such a colour region may have its own local boundary, as would be required by the popular meaning of whiteness, or be extended to cover the whole colour diagram, if a meaning parallel to colorimetric desaturation be adopted. General agreement on this point and on corresponding graded requirements in regard to accuracy under various conditions is very desirable.

With reference to the Blancometer, it may be generally observed that as one may safely assume that the near-whites met with in practice are, almost without exception, of restricted selectivity approximately characteristic of their colour, the accuracy of colour measurement by such means, employing a colorimetrically corrected photocell-filter combination, must necessarily be sufficient for any given purpose within some definite colour region, and I am pleased to note Dr. Harrison's remarks in this connection from the practical standpoint. The work which is being done by him and others in investigating and evaluating the traditional generalizations and practical concepts employed in ancient arts and trades is commendable, and as unknown phenomena may lie obscurely hidden in the application of industrial processes, such work may also be fruitful by bringing these to light. Such is the case apparently with the curious phenomenon relating to subjective-brightness comparisons which he describes and which, if genuine, may be of associative origin. Regarding these tests, I should like to ask him whether any special illuminant was used and the possible effect investigated of variations of colour temperature upon the results of the comparisons; also whether the effects of gloss were eliminated.

If this phenomenon is purely subjective and allied to colour temperature, it may have applications in the projection of colour films within the relatively broad colour tolerances there possible, through the use of a filter or mirror which would suitably re-adjust the implied colour temperature of the object " imaged " on the screen, to produce an apparent brightness increase.

Mr. H. G. W. Harding. As regards the paper by Mr. Holmes, the relative merits of magnesium oxide and titanium oxide for a standard white have been mentioned by Mr. Holmes and Dr. Jolley. In most colour measurements it is necessary to have a standard white, the reproducibility of which is above suspicion. At the National Physical Laboratory we have used as a standard a polished silver plate on to which magnesium oxide has been deposited to a thickness of about 2 mm. A few measurements that I made some years ago indicated that although the colour and brightness of such a deposit varied appreciably between thicknesses of 0·01 mm. and 1·0 mm., the changes from 1·5 mm. to 2·0 mm. were very small. Deposits thicker than these are difficult to obtain since they fall away from the silver surface.

Dr. Harrison, in his paper, mentions the suitability of the Hardy spectrophotometer that is available in America for the work in which he is concerned. There appear to be no instruments of this type available here, but there are non-automatic photo-electric instruments which give satisfactory results. The time taken for a spectral reflection curve defined by forty points using the instrument at the National Physical Laboratory is about an hour. An instrument of this type may be of use.

Mr. **Applebee.** I wish to point out that, in addition to the factor mentioned (namely, the " colour " of the white screen and the " colour " of the various light sources), the resultant reflected light from the maintained lighting (i.e. the lighting that has to remain alight during the showing of the picture) and also from the tobacco-smoke-charged picture beam should be considered. Both these cause stray light to be reflected on to the screen, and this stray light is again likely to vary in each building owing to the dominating colours of the decorative motive of the walls. It may therefore give to the screen an orange tinge, or, in another cinema, possibly painted in the light greens, a blue tinge; add to this the " nicotine " colouring of the smokers, together with the insistence of certain licensing authorities in some parts of the country on a higher intensity of secondary and maintained lighting, and you have a most complicated problem when it comes to standardization.

Dr. W. D. **Wright.** I should like to ask Mr. Heys Hallett how far the standardization of the colour of cinema screens in this country would be affected by any corresponding standardization abroad, especially in the United States. If different standards were adopted in the two countries would it be possible to secure adequate compensation by modifying the printing conditions of colour positives?

I should also like to ask Mr. Heys Hallett whether he considers it possible and worth while to express his colour measurements in terms of the standard C.I.E. system or whether he regards the cinema problem as so specialized that an arbitrary system of measurement is to be preferred?

Mr. R. **Donaldson.** One of the problems that Dr. Harrison has discussed is very accurate colour measurement over a small part of the colour field, viz., the yellowish-white region. The trichromatic colorimeter is designed to measure all colours, both saturated and unsaturated, and to do this it must have a small photometric field and saturated primaries. The colorimeter can, however, be easily adapted to cover a small portion of the colour field with increased accuracy.

Greater precision of setting can be obtained by using the large Lummer-Brodhun contrast field instead of the normal 2° field. With the large field the eye works under very sensitive conditions, and any colour difference visible to ordinary viewing becomes more apparent in the Lummer-Brodhun field. The personal error due to the colour-vision characteristics of the observer may be reduced in two ways. Pale unsaturated filters can be substituted for the usual deep red, green and blue filters supplied with the instrument, so that big differences in spectral-energy distribution between the two halves of the photometric field are avoided. Alternatively it may be sufficient to choose a comparison standard close to the colour being measured. If the magnesium oxide is too far away a piece of calibrated matt opal glass may be suitable as a standard.

AUTHORS' replies to discussion.

Mr. J. G. HOLMES. In reply to Dr. Jolley, it has been my experience that different consignments of titania may have different colours, ranging from white to pale yellow, which cannot be correlated with any traces of impurities. This difference is probably attributable to the technique of manufacture, and renders the oxide unsuitable as a general reference standard. Titania forms a good undercoating for magnesium oxide and gives an opaque white of high reflection factor with only a thin coating of magnesium oxide. A better absolute standard is undoubtedly given by the silver-plate method described by Dr. Harding.

In reply to Mr. Perry's comments on the noun " whiteness ", there is no fundamental difference between the whiteness and the redness or yellowness of a colour. I have recently been able to find the limits of the area of the colour chart within which coloured light sources are called white, red, yellow, etc., and small areas of this type have been used for several years as a definition of a colour in standard specifications. When considering coloured materials, however, whiteness has a different meaning from redness or yellowness, but the popular meaning of the word is so deeply established that it would be difficult to give it any quantitative significance which would be generally acceptable.

Dr. V. G. W. HARRISON. In reply to Mr. Donaldson, we have had our Donaldson colorimeter fitted with Lummer-Brodhun contrast field, which can be used alternatively to the normal 2° field. With the large field there is undoubtedly a considerable gain in sensitivity, but the reproducibility of readings for different observers and even for the same observer on different days remains disappointing. Mere sensitivity without stability and reproducibility is useless in industrial work. I have not tried in the colorimeter the desaturated filters which Mr. Donaldson mentions, but we have had some experiments made with the Lovibond tintometer, which makes use of very desaturated filters, although the colour mixture here is subtractive and not additive. The errors with the Lovibond instrument were of the same order as those observed with the Donaldson colorimeter. While I do not claim to have said the last word on this subject, from the evidence available to me I consider that it is unlikely that visual colorimeters can be improved to such an extent that they will be able to satisfy the very exacting demands of commercial work.

In reply to Mr. Harding, the question whether a non-recording spectrophotometer is suitable for measuring the colours of papers depends on how many papers have to be measured during the day. For research purposes, where only two or three curves are required per day, the type of instrument in use at the National Physical Laboratory should be quite satisfactory; but for commercial work a non-recording instrument will be much too slow. It is of interest to note that Hardy developed his recording spectrophotometer precisely because he found available non-recording types much too slow for normal routine work.

As regards the question put by Mr. Perry, the tests on " visual brightness " were made in diffused daylight, direct sunlight being avoided. The experiments were carried out in the late morning and early afternoon of a single day, and there were no marked changes of weather during the period of the tests. I should not care to forecast what would be the effect of large changes in colour temperature on the results. The samples were laboratory sheets with a rough matt surface having no perceptible gloss.

I should like to point out that I started tests as an avowed sceptic and finished by " discovering " something which had already been known for many years in the paper trade; I have thus good reason to believe that the results are genuine.

Mr. **Heys Hallett**. The answer to Dr. Wright's first question is that projection illuminants are much the same in all countries, and it is therefore most improbable that a country would set up a standard that would cause any anxiety to the printers. It is probably desirable to avoid having to make special prints for each country, and an International Standard should therefore be aimed at. I believe that this country and the U.S.A. are the only ones having colour-printing laboratories. No other country has, as yet, put forward any proposals.

With regard to his second question, it is desirable, within the cinema industry, to express the results in a manner which obviates the necessity for any knowledge of photometry, and particularly to eliminate the possibility of confusion between the common procedure of measuring the incident light and the newer one of measuring the brightness. It is for these reasons that we developed the idea of specific brightness and specific colour.

I think it is most important that the standards, when laid down, should be linked to the standard C.I.E. system in order that they may be reproducible, and to allow instruments such as the Morganite Reflectometer to be properly calibrated. I hope, therefore, that the standards, when adopted, will be laid down in terms of the C.I.E. system, but that expression as specific brightness will continue to be used in kinemas. I hope shortly to express our present datum in terms of the C.I.E. system.

In reply to Mr. Applebee, I agree that maintained lighting, light scattered from the projection beam and the colour of the decorations are not without effect, though usually small enough to be negligible. The maintained lighting installation should be so designed that none falls on the screen, but if the measurements are made with the maintained lighting switched on, any direct effect which they have will be taken into account.

These three causes may have more effect on the picture by contrast than by direct interference, but this is a complex subject, outside the scope of the present paper.

THE MEASUREMENT OF IMPULSE VOLTAGES BY MEANS OF SMALL SPHERE GAPS

By D. E. M. GARFITT, B.Sc. Tech.,
Metropolitan-Vickers Electrical Co., Ltd.

Communicated by Dr. A. P. M. Fleming, 14 *November* 1941

ABSTRACT. The erratic behaviour of small sphere gaps under impulse voltages may be explained by the comparative scarcity of ions in the gap. Methods are described by which this deficiency of ions may be remedied and greater consistency of breakdown be obtained. Suggestions are made for the use of the 0·5-inch diameter sphere gap as a standard for impulse-voltage measurement when irradiated by light from a spark discharge.

§ 1. INTRODUCTION

This investigation arose from an attempt to measure small impulse voltages of very short duration by means of a 0·5-inch-diameter sphere gap. During the measurements it was observed that for the same applied voltage, wide variation was obtained in critical sparking distance and in brightness of spark. Means had therefore to be sought for improving the consistency of the gap, to enable useful calibrations to be made. This could be accomplished by supplying the gap with ionization from an external source, as is described in §§ 3 and 4 below.

§ 2. PROPERTIES OF THE GAP

When impulse voltages are applied to a fixed gap, it is found that between the maximum voltage at which sparkover never occurs and the minimum voltage at which sparkover is always obtained there is an appreciable range in which sparkover occurs intermittently. The impulse breakdown voltage of a gap has been arbitrarily defined in B.S.S. 923/1940 as that voltage for which 50% of the applications will cause breakdown.*

---

* Afterwards abbreviated to B.D.

The range over which variation is obtained is expressible in terms of "dispersion", which may be defined arbitrarily as:

$$\frac{\text{Voltage required to give 90\% B.D.} - \text{Voltage required to give 10\% B.D.}}{\text{Voltage required to give 50\% B.D.}}$$

(The choice of 90% and 10% is purely one of convenience which has proved to be of good service.)

With small spheres at low spacings, the impulse breakdown voltage may be appreciably higher than the corresponding figure for a constant unidirectional voltage. The *impulse ratio* for any gap is defined as:

$$\frac{\text{Impulse voltage to give 50\% B.D.}}{\text{Constant unidirectional breakdown voltage for same gap}}.$$

Impulse ratio and dispersion are closely related to the time lag of the gap—i.e. to the interval elapsing between the attainment by the impulse of the static breakdown voltage of the gap and the actual breakdown of the gap—which is known to vary with wave-shape. The gap calibration might, therefore, be expected also to depend upon wave-shape.

§ 3. EXPERIMENTAL CALIBRATION OF THE SPARK GAP

The gap under test consisted of two 0·5-inch diameter spheres whose sparking distance could be adjusted to ±0·3 mil by means of a micrometer scale. The test voltage was obtained from a single-stage impulse generator of which the circuit is shown in figure 1. The charging voltage V could be measured to within ±0·03 kv. by a resistance voltmeter which consisted of a wire-wound oil-immersed resistance of 18 mΩ in series with a sub-standard milliammeter. The resultant accuracy of measurement was thus in the neighbourhood of ± 6% for a gap of 10 mils, and ± 0·6% for a gap of 100 mils. The impulse voltage was deduced from the charging voltage by computing the

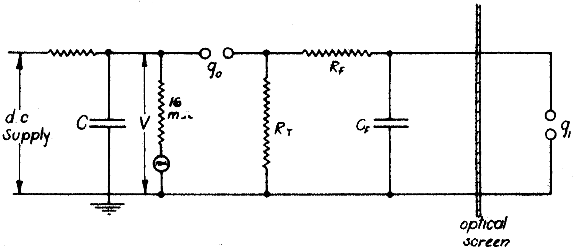


Figure 1. Circuit employed for impulse calibration of sphere gaps.

Circuit constants

C_F μF	R_F ohms	C μF	R_T* ohms	Wave-shape μ sec.	Efficiency of generator %
0·0005	100	0·25	25·6	0·15/4·5	96
0·0005	100	0·25	580	0·15/100	100
0·0005	420	0·25	25·6	0·63/4·5	89
0·0005	420	0·25	580	0·63/100	99

efficiency of the impulse generator from the circuit constants, by the method given in Appendix 2. As no account has been taken of the stray inductance and capacitance of the circuit, the absolute accuracy of the calibration will be somewhat less than the value given above, although comparison of curves taken with the same circuit constants will be unaffected.

Readings were normally taken with negative impulse voltage. A number of readings were repeated with positive impulse voltage at sparking distances up to 100 mils, but any difference due to the change of polarity was less than the probable error of measurement. This is supported by work on larger sphere-gaps (Meador, 1934; Bellaschi and McAuley, 1934), which has shown that polarity difference is first detectable at a spacing of 20% to 25% of sphere diameter.

It was necessary to ensure that the gap g_1 received no radiation from unwanted sources of light, particularly from the gap g_0 of the generator, except where radiation from this source was specified in the calibration. Reflection of light from varnished or polished surfaces had also to be avoided by suitable screening. The spheres were cleaned frequently with metal polish and ether; cleaning appeared to have little effect upon the behaviour of the gap, except in the case of magnesium spheres, when the consistency was impaired by the oxide coating which formed rapidly under sparking. Dust on the cathode might be expected to affect results dependent upon photo-ionization at the metal surface, but seems to be relatively unimportant, probably because in a dusty atmosphere a very large proportion of the precipitation occurs upon the positively charged surfaces.

To enable impulse ratios to be calculated, a calibration of a brass sphere-gap was made up to 10 kv. with constant unidirectional voltage. Radiation from 0·5 mg. of radium was used to give consistency as described in § 4 (*a*) I below. No standard calibration tables include the 0·5-inch diameter sphere gap, but a calibration given by Klemm (1923) under radiation from a mercury arc will be seen from table 1 to be in close agreement with that obtained by the author.

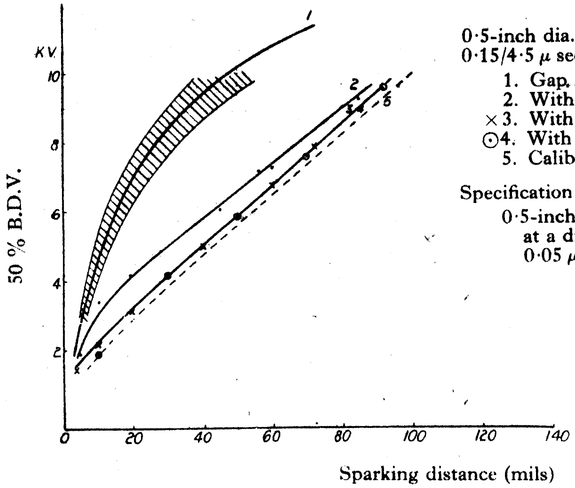
Table 1

Separation of spheres (mils)	Static B.D.V. with 0·5 mg. Ra. (kv.)	B.D.V. obtained from curve drawn from Klemm's data (kv.)
10	$1{\cdot}8_1$	$1{\cdot}8_3$
20	$2{\cdot}9_0$	$2{\cdot}9_0$
30	$3{\cdot}8_3$	$3{\cdot}8_4$
40	$4{\cdot}7_3$	$4{\cdot}7_2$
50	$5{\cdot}6_2$	$5{\cdot}6_1$
60	$6{\cdot}5_8$	$6{\cdot}4_8$
80	$8{\cdot}1_7$	$8{\cdot}1_8$
100	$9{\cdot}6_7$	$9{\cdot}8_8$

It should be noted that in all calibrations given here, the voltages have been reduced to their equivalent values at 20° c. and 760 mm. of mercury.

$$V_{20,760} = V_{t\cdot p}/\delta \text{ where } \delta = 0{\cdot}386p/(273+t).$$

Figure 2 contains calibration curves taken under various conditions of ionization with an impulse of 0·15/4·5 microseconds, and in figure 3 are shown



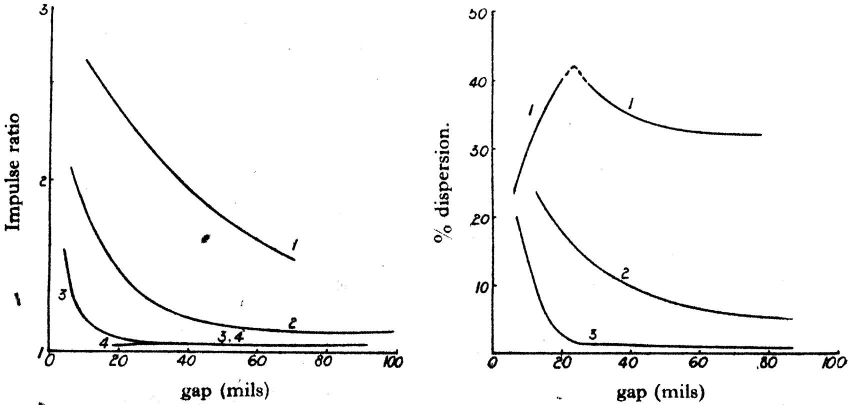
0·5-inch dia. brass spheres.
0·15/4·5 μ sec. negative impulse voltage.

1. Gap not irradiated.
2. With 0·5 mg radium.
×3. With polonium-coated sphere.
⊙4. With radiation from another spark gap.
5. Calibration on d.c.

Specification of auxiliary spark :—

0·5-inch brass sphere gap, spacing 40 mils, at a distance of 12″. Connected across 0·05 μF. without series resistance.

Figure 2. Calibration of 0·5-inch dia. sphere gap with d.c. and impulse voltages. Brass spheres. 0·15/4·5 μ sec. negative impulse voltage.



(*a*) (*b*)

Figure 3. (*a*) Impulse ratio and (*b*) dispersion of breakdown voltages.

1. Gap not irradiated.
2. With 0·5 mg. radium.
3. With polonium-coated sphere.
4. With radiation from another spark gap.

the corresponding values of impulse-ratio and dispersion. These results are discussed fully below. Curve 1 was plotted for brass spheres screened from radiation. The hatched area indicates the envelope of points obtained on different occasions. Since results are not repeatable within wide limits, such a calibration is clearly useless for purposes of accurate measurement. If, however, the gap receives a large supply of ions at the time of application of the impulse, much greater consistency is obtained. Ionization may be produced in the gap by the action of

(*a*) the products of disintegration of radioactive substances.
(*b*) ultra-violet radiation from an arc or spark discharge.

§ 4. CALIBRATION OF A GAP SUPPLIED WITH IONIZATION

(*a*) *The use of radioactive substances*

(I) Radium is an ionizing agent frequently employed. In the method devised by Van Cauwenberghe (1930) the radium salt is contained in a small capsule which is inserted in one sphere. Curve 2, figure 2, was obtained with 0·51 mg. of radium element in a container of monel-metal, 0·2 cm. thick, inserted in a brass sphere which had been drilled out to give a wall thickness of 1 mm. of metal between the capsule and the gap. Absorption in the metal would eliminate all but γ radiation, which may be considered as solely responsible for the improvement in the gap. The capsule was found to be equally effective in either anode or cathode. Though the curve shows considerable improvement in the performance of the gap, individual points deviated by 4 % from the mean curve. At spacings of 50 mils or less, dispersion was of the order of 10 % and determination of the 50 % sparkover value required an impracticably lengthy series of readings.

Figure 4 (*a*) shows results obtained with radium as ionizing agent for four different wave-shapes. It will be seen that dispersion and impulse-ratio increase as the impulse tail becomes shorter; this might be expected, as with a longer tail, greater time-lags may be attained and over-voltages will therefore tend to be reduced. Dispersion and impulse ratio are also increased at lower spacings when the wave-front becomes longer. This is possibly because the more slowly rising voltage tends to sweep ions out of the gap before the crest is reached.

(II) Ionization may be provided by active material on the surface of the electrodes (Dillon, 1940). Curve 3, figure 2, was obtained with silver-plated brass spheres, one of which had been coated with a thin film of polonium (for note on intensity of α-particle and γ-ray ionization see Appendix 2). In this case, the emission consisted of α particles with a range of 3 cm. Silver-plated spheres were tested without irradiation and were found to behave in the same erratic manner as brass, therefore it was assumed that the presence of the silver had no effect upon breakdown. It will be seen from the curve that for gaps of 30 mils or more (i.e. at voltages higher than 4 kv.) the dispersion is less than

1 % and the impulse ratio is nearly unity. Results obtained with the polonium coating for other wave-shapes are shown in figure 4 (*b*). The effect of wave-shape upon breakdown voltage, though still evident at lower spacings, is much smaller than when a less powerful ionizing agent is used. It would seem, therefore, that this method of irradiation is suitable for the measurement of impulse voltages above about 4 kv. It should, however, be noted that the polonium coating has a half-life period of 136 days, and will require fairly frequent renewal, though no deterioration of its influence on the calibration was observed 53 days after its preparation. It is also possible that heating of the spheres due to passage of sparks will cause the polonium to volatilize and thus further reduce its useful life.

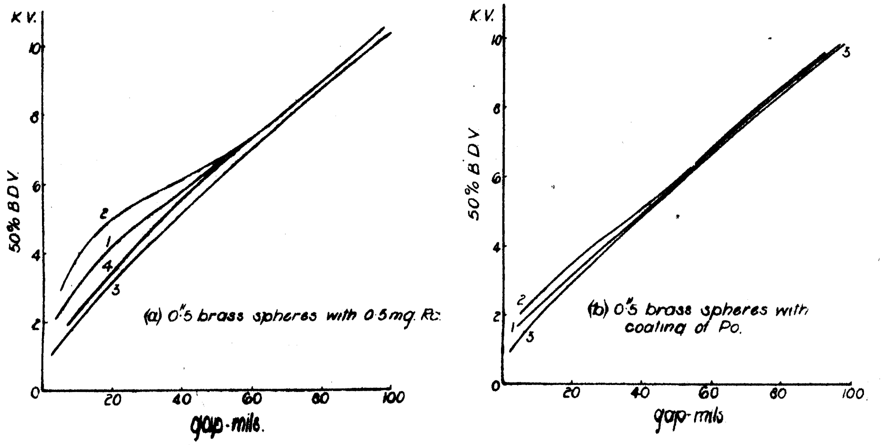


Figure 4. Comparison of B.D. of 0·5-inch sphere gap under impulse voltages of different shapes.

1. 0·15/4·5 μs impulse voltage.
2. 0·63/4·5 ,, ,, ,,
3. 0·15/100 ,, ,, ,,
4. 0·63/100 ,, ,, ,,

(*b*) *The use of ultra-violet radiation*

(I) Radiation from a mercury arc in quartz is an ionizing agent that has been adopted by Van Cauwenberghe (1930) and others (Nord, 1935). Results were obtained by the author with a 3-inch arc at a distance of 2 ft. 6 in. from the gap, but as this did not appear to be any more effective than 0·5 mg. of radium and was less convenient mechanically, the method was not pursued further.

(II) The spark discharge itself is known to be a source of considerable ultra-violet radiation, and the effect of irradiation of one spark gap by light from another has been studied by several workers (e.g. Brinkmann, 1939), but the method has not yet been applied to the production of a standard of impulse-voltage measurement.

Two alternatives appear:—

(i) Light from the gap g_0 of the impulse generator may be allowed to fall upon the impulse spark gap, or

(ii) The impulse gap may be illuminated by an independent third gap, and provision made for this gap to operate at the correct instant with respect to the arrival of the impulse.

These two methods are illustrated in figure 7.

The radiation from a spark varies very much in effectiveness, depending upon the voltage across the gap, the values of capacitance and resistance in the discharge circuit, and probably upon the metal of the electrodes.

To obtain repeatable results, therefore, it is essential to specify the circuit of the illuminating spark in detail.

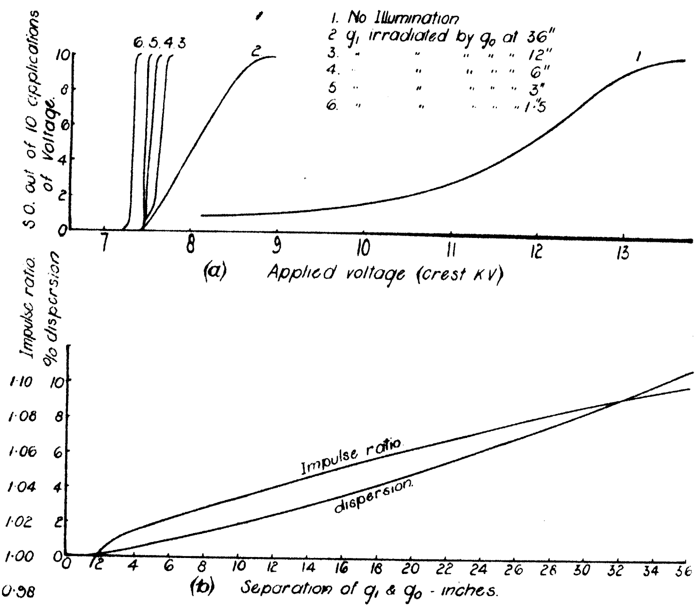


Figure 5. B.D. of 0·5-inch dia. sphere gap under illumination from another spark. Variation of intensity of illumination by adjustment of distance between sparks.

0·15 μsec. negative impulse voltage $g_1 = 70$ mils.

(i) The curves of figures 5 and 6 were obtained when the impulse-generator gap was used also as the source of illumination. Since the discharge circuit of this gap includes R_T, comparison of results obtained with different waves is not permissible. Figure 5 illustrates the effect of variation of the distance between the two gaps. As the intensity of illumination increases, the width of the dispersion band is reduced until it is negligible, and the impulse ratio approaches unity. Further increase in illumination shifts the curve as a whole towards the origin, thus causing the 50 % breakdown value to decrease more rapidly, until it attains a value lower than that of the normal d.c. breakdown. In this case the impulse ratio, is apparently less than unity, but this is only a matter of definition of impulse ratio, as illumination of the same intensity would probably reduce the static breakdown voltage to a value equal to or less than the impulse breakdown voltage.

The effectiveness of the spark illumination was reduced appreciably when a barrier of frosted quartz, 0·04 inch thick, was placed between the gaps. When the gaps were 4 inches apart, the quartz increased the breakdown voltage by 1 %, and the dispersion rose from 0·4 % to 0·8 %. With the gaps at a distance of 2 inches, the corresponding increases were 4 % in breakdown voltage and 30 % in dispersion (from 0·3 % to 0·45 %).

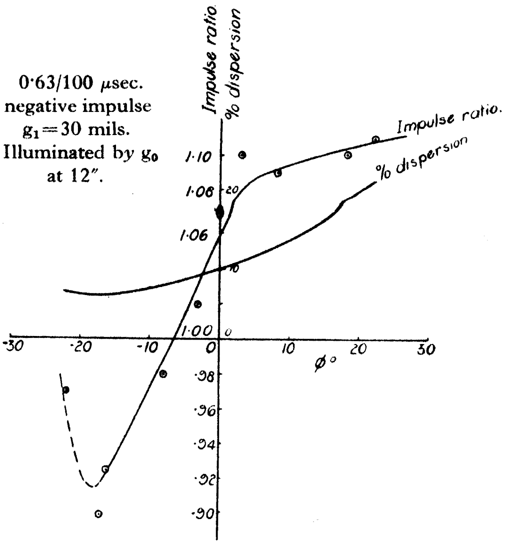


Figure 6. B.D. of 0·5-inch sphere gap under illumination from another spark. Variation of illumination on the cathode.

Figure 6 shows the effect of increasing the proportion of illumination falling on one electrode, while keeping the total illumination constant. The gaps were set up as shown in figure 7 (*a*). The impulse gap was at a fixed distance from the illuminating gap (measured between gap centres), but could be rotated through an angle ϕ, which was taken as positive when the cathode was displaced towards g_0, i.e. when the anode received more illumination. It will be seen from figure 6 that the light is most effective when the angle ϕ is of the order of $-18°$; this is approximately the angle Φ at which the path of the light is along the common tangent to the two spheres, when the cathode receives maximum illumination. These results support the view that the effectiveness of the light is due to photo-emission from the cathode rather than to volume ionization in the gap.

(ii) When a third gap is used as illuminating source, a very intense spark may be obtained by supplying the gap from a large condenser, without series resistance. Direct comparison of results with different wave-shapes is permissible with this method.

A calibration obtained with illumination from a third gap is shown in curve 4, figure 2. The illuminating gap consisted of brass spheres, spaced at 40 mils (corresponding to 4·7 kv.) and supplied from a condenser of 0·05 μF., at a distance of 12 inches from g_1. To ensure that the impulse gap was illuminated at the instant of arrival of the impulse, it was necessary to allow light from the third gap to trip the impulse generator. This was secured by the arrangement shown in figure 7 (*b*). The voltage *V* across g_0 was held at a value below the normal breakdown voltage of the gap, but above the breakdown value obtained with intense spark illumination. Passage of a spark at the illuminating gap then caused the generator to trip. To obtain sufficient intensity of illumination at g_0 it was necessary to set it within 6 inches of the illuminating gap. Pick-up in the generator circuit was avoided by surrounding the illuminating gap by an earthed screen, complete except for two small apertures to permit illumination of g_0 and g_1. g_1 was carefully screened from direct light from g_0, though this was probably unimportant, as the spark at g_0 was much less intense than the illuminating

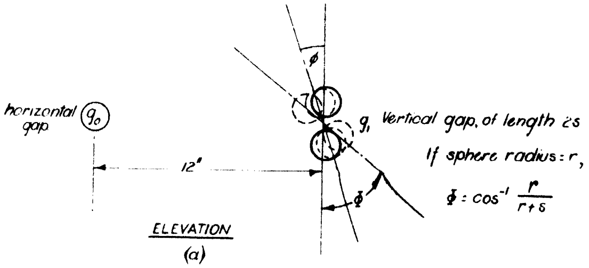

horizontal gap
g_0
12"
ϕ
g_1 Vertical gap, of length δ
If sphere radius = r,
$\phi = \cos^{-1} \frac{r}{r+\delta}$
ϕ
ELEVATION
(a)

Figure 7 *a*. Illumination of the impulse gap by light from the impulse generator.

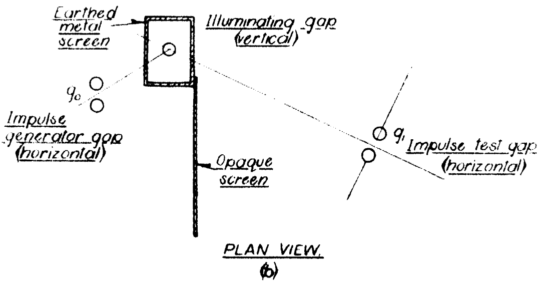

Earthed metal screen
Illuminating gap (vertical)
g_0
Impulse generator gap (horizontal)
Opaque screen
g_1
Impulse test gap (horizontal)
PLAN VIEW.
(b)

Figure 7 *b*. Illumination of the impulse gap and tripping of the impulse generator by light from a third gap.

spark. Both electrodes of the impulse gap were equally illuminated, but the influence of small variations in the angle and the distance between the gaps seemed less critical than when the impulse generator gap was used as illuminating agent. It is likely that the relative times of occurrence of maximum intensity of illumination and impulse crest voltage differ in the two methods of spark illumination described here, and that the interval between the two maxima is of importance in determining the influence of the spark.

From figures 2 and 3 it will be seen that the calibration obtained by this method at higher spacings is almost the same as that obtained with polonium, and at lower spacings is of much higher accuracy. Dispersion was negligible over the whole range tested, and the impulse ratio was less than 1·02.

§ 5. SPHERES OF MAGNESIUM ALLOY

The use of electrodes of a metal of low work-function, from which photoelectrons may be liberated under the incidence of normal daylight, was considered as a means of obtaining consistent breakdown. Magnesium has a work-function of 2·7 ev., whereas the values for copper and zinc, of which brass is composed, are 4·0 ev. and 3·4 ev. respectively (these are the values given in Kaye and Laby's *Tables of Physical and Chemical Constants*).

Spheres of magnesium were tested in daylight, but without other radiation. At a gap of 50 mils, dispersions of 25 % and 35 % were obtained with wave tails of 100 and 4·5 microseconds respectively and the impulse ratio was of the order of 1·6 in both cases. Accurate measurement was rendered more difficult by the rapid formation of an oxide film, which, unless cleaning was frequent, rendered the gap almost as inconsistent as one of brass. With brass spheres, the corresponding figures were for long and short tails 33 % and 35 % dispersion, and 1·7 and 1·8 impulse ratio. It seems, therefore, that there is little advantage to be gained by the use of magnesium electrodes.

§ 6. CONCLUSIONS

(1) Small sphere gaps without irradiation are unsuitable for the accurate measurement of impulse voltages of short duration. No advantage is gained by the use of spheres of magnesium.

(2) With 0·5 mg. of radium in the gap, consistency is considerably improved, but is still insufficient to permit of measurements accurate to more than ± 5 % at 5 kv. An improvement of the same order is obtained under radiation from a mercury arc in quartz.

(3) A satisfactory calibration, repeatable to within 1 %, is obtained at voltages of 4 kv. upwards when ionization is obtained from polonium deposited on one of the electrodes. The comparatively short life of the active film is, however, a disadvantage.

(4) If radiation from a spark discharge is employed as an ionizing agent, the consistency of breakdown may be so high that the resultant accuracy of

measurement is governed entirely by the accuracy in setting the measuring gap, provided that the conditions under which the illuminating spark is obtained are reproduced as closely as possible whenever the calibration is used.

§ 7. ACKNOWLEDGEMENTS

The author is indebted to Dr. T. E. Allibone for advice and criticism, and to Mr. F. S. Edwards and Mr. J. M. Meek for stimulating discussion. Thanks should also be expressed to Dr. N. Feather (Cavendish Laboratory, Cambridge) for the preparation of the polonium surface, and to Dr. A. P. M. Fleming, C.B.E., Director and Manager of Research and Education Departments, Metropolitan-Vickers Electrical Co., Ltd., for permission to publish this paper.

REFERENCES

Bellaschi, P. L. and McAuley, P. H., 1934. *Elect. J.* **31**, 228.

Brinkmann, C., 1939. *Arch. Elektrotech.* **33**, 1, 121.

Van Cauwenberghe, R. and Marchal, G., 1930. *Rev. Gén. Elect.* **27**, 331.

Dillon, J. H., 1940. *J. Appl. Phys.* **11**, 291.

Klemm, A., 1923. *Arch. Elektrotech.* **12**, 553.

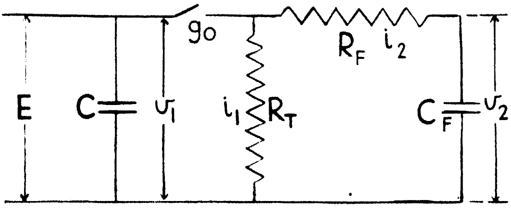
Meador, J. R., 1934. *J. Amer. Instn Elect. Engrs*, **53**, 942.

Nord, G. L., 1935. *J. Amer. Instn Elect. Engrs*, **54**, 955.

APPENDIX 1

The estimation of correction factors for the determination of impulse voltages from a knowledge of the impulse-generator charging voltage

The impulse generator may be represented schematically as follows:—



If g_0 is closed at time $t=0$, then v_2 is subsequently

$$v_2=\frac{1}{C_F}\int i_2 dt \qquad \text{......(1)}$$

Also

$$R_F i_2+v_2=v_1, \qquad \text{......(2)}$$

$$i_1=v_1/R_T \qquad \text{......(3)}$$

and

$$E=\frac{1}{C}\int(i_1+i_2)dt+v_1. \qquad \text{......(4)}$$

Therefore

$$E=v_1+\frac{1}{C}\left\{\frac{\int v_1 dt}{R_T}+v_2 C_F\right\} \qquad \text{......(5)}$$

and

$$v_1=v_2+R_F C_F dv_2/dt. \qquad \text{......(6)}$$

Now, from (1) $$0=Dv_2+R_FC_FD^2v_2+\frac{1}{C}\left(\frac{v_2}{R_T}+\frac{R_FC_F}{R_T}Dv_1+C_FD_Fv_2\right),$$

i.e. $$\left[R_FC_FD^2+\left(1+\frac{C_F}{C}+\frac{R_FC_F}{R_TC}\right)D+\frac{1}{R_TC}\right]v_2=0.$$

Substituting $$T_1=CR_T,\ T_2=C_FR_F,\ X=C_F/C,$$

we have $$\left[T_2D^2+\left(1+X+\frac{T_2}{T_1}\right)D+\frac{1}{T_1}\right]v_2=0.$$

This equation has roots

$$\alpha=\frac{-\left(1+X+\frac{T_2}{T_1}\right)}{2T_2}\left[1\pm\sqrt{\left(1-\frac{4T_2}{T_1\left(1+X+\frac{T_2}{T_1}\right)^2}\right)}\right].$$

The solution is of the form $v_2=Ae^{-\alpha_1t}+Be^{-\alpha_2t}$.

But at $t=0$, $v_2=0$; therefore

$$v_2=A(e^{-\alpha_1t}-e^{-\alpha t_2}).$$

Time to crest t_1:— $$Dv_2=A(-\alpha_1e^{-\alpha_1t}+\alpha_2e^{-\alpha_2t})$$

$$=0 \text{ when } (\alpha_2-\alpha_1)t=\log_e\frac{\alpha_2}{\alpha_1},$$

i.e., when $$t_1=\frac{\log_e\alpha_2/\alpha_1}{\alpha_2-\alpha_1}.$$

At $t=0$, C_F behaves as a short-circuit,

and $$Dv_2=i_2/C_F=E/T_2,$$

i.e. $$A(-\alpha_1+\alpha_2)=E/T_2,$$

and for unit voltage on C, $$A=\frac{1}{T_2(\alpha_2-\alpha_1)}.$$

Therefore the impulse voltage $v_2=\eta E$, where the correction factor η is given by

$$\eta=A(e^{-\alpha_1t_1}-e^{-\alpha_2t_1})=\frac{e^{-\alpha_1t_1}-e^{-\alpha_2t_1}}{T_2(\alpha_2-\alpha_1)}.$$

APPENDIX 2

The silver sphere was coated with a layer of polonium by immersion in a hot solution of RaD, E and F in ½N HCl for about an hour. Almost half the sphere surface was activated, and the available α-ray intensity was estimated to be $0{\cdot}3\times10^7$ particles/cm²/sec. leaving the sphere. This results in an intensity of ionization of 8×10^{10} ion-pairs/cm³/sec. near the sphere surface. The α-ray intensity from the radium capsule is estimated to be about 4×10^7 ion-pairs/cm³/sec. near the sphere surface, so the intensities differ by a factor 10^3 to 10^4.

THE NATURE OF TEMPERATURE

By W. E. BENHAM

MS. received 6 *September* 1941

ABSTRACT. There is at first sight some difficulty in reconciling temperature as energy per unit mass with the conception of temperature obtainable from radiation theory. Evidence is adduced to support the view that the ultimate significance of temperature is that it is measured by the thinness of a pulse of electromagnetic radiation. Some preliminary remarks on a new theory of radiation (which requires that central orbits shall be non-radiating when circular) are used to support a dimensional treatment in which the energy density of radiation is a function of the mass (rather than of the charge) of an electron and of the absolute temperature. The actual dimensions of temperature appear to be 1/L, and this is reconciled with energy per unit mass by attributing dimensions to "Newton's constant" N. It turns out that this and the gravitational constant have the dimensions

$$[N]=\mathrm{T}^2/\mathrm{L}^3,$$
$$[G]=1/\mathrm{M}.$$

The Wiedemann-Franz ratio has dimensions M/L^2, and a number of thermal quantities are listed as to their dimensions on the basis proposed. A quantity known to be an adiabatic invariant has the dimensions T/L and the velocity involved is thought to be that of electromagnetic waves.

§ 1. INTRODUCTION

Lord Kelvin's proposal to eliminate mass as a fundamental unit was criticized in a recent discussion,* the keynote of which is the retention of what will be termed "Newton's constant" in the mass-acceleration law

$$F=Nma. \qquad \ldots\ldots(1)$$

While N may be taken for convenience as of unit magnitude, this makes it too easy to forget the dimensions of N, which should therefore be retained until valid reasons arise for doing otherwise.

We shall now take temperature as energy per unit mass, understandable from the kinetic theory and from calorimetry with heat as a form of energy. If we do not take N as dimensionless, we no longer find that temperature has the dimensions [velocity]^2, or, writing θ for absolute temperature, T being wanted for the dimensions of time, Θ being the dimensions of temperature, M, L those of mass, length, respectively, we do not find

$$\Theta=\mathrm{L}^2/\mathrm{T}^2 \qquad \ldots\ldots(2)$$

but

$$\Theta=[N]\mathrm{L}^2/\mathrm{T}^2. \qquad \ldots\ldots(3)$$

---

* *Proc. Phys. Soc.* 1940, **52**, 608.

We shall show that the retention of N, that is, the retention of generality so dear to mathematical physicists, except, apparently, when the subject is that of " dimensions ", is of assistance in clearing up an apparent discrepancy in the dimensions of temperature as derived above, on the one hand, and as obtained either from the classical theory of radiation by collisions or from the principle of adiabatic invariance, used in the quantum theory of radiation, on the other hand. Let us first see where Kelvin's logic will lead us. He says, in effect, that if we measure in units for which G, the gravitational constant, as well as N, is dimensionless, the proper measure of a force is the fourth power of a linear velocity, while mass per unit volume is measured by the square of an angular velocity, or

$$F = L^4/T^4 \ ; \quad M = L^3/T^2. \qquad \ldots\ldots(4)$$

Now, whilst it is one thing to take some quantity or other as the proper measure of a physical entity, it is quite another to identify the dimensions of the entity with those of the quantity, as is done under (4). Thus, L^4/T^4 has been taken as the physical dimensions of force. Once we start doing this sort of thing, there is no knowing where we shall end up. We might just as well say that because the energy density of black-body radiation varies as the fourth power of the absolute temperature and is independent of the material giving rise to the radiation, therefore Θ^4 will represent the physical dimensions of either energy per cubic centimetre or energy per square cm. per sec., which leads to two possibilities:—

$$\Theta^4 = M/T^2L \ ; \quad \Theta^4 = M/T^3, \qquad \ldots\ldots(5)$$

neither of which is reconcilable with (2), to which Kelvin is logically committed, or, saving fractional indices for Θ, with the second of (4).

§ 2. DISCUSSION OF STEFAN'S LAW

Let us begin by calling attention to the formula of classical electromagnetic theory for the energy per unit volume of a stream of radiant energy at a place where the dielectric constant * is k and the electric vector has the value E, namely,

$$\epsilon = kE^2/4\pi. \qquad \ldots\ldots(6)$$

Equation (6) is valid in any set of units provided k is appropriately expressed, but we shall regard it as being in electrostatic units. The energy due to the magnetic vector of the field is of course included in (6), being equal in magnitude to half the total.

Now the electric vector arises, in the last analysis, from electric charges. In the study of (6) from the dimensional standpoint we need only concern ourselves with this qualitative statement. Replacing E in terms of charges of dimensions Q, but of unknown or unspecified magnitude or location, we have the dimensional equation (where K denotes the dimensions of k)

$$[\epsilon] = Q^2/KL^4. \qquad \ldots\ldots(7)$$

---

* Regarded as including any possible " force constant ".

We next point out that the above dimensional relation must hold good in the particular case of black-body radiation, for which ϵ varies as θ^4. Now, in the classical theory of radiation by collisions, it turns out that the time occupied by a collision of an electron with a molecule is inversely proportional to the absolute temperature. To quote Sir J. J. Thomson (1907a): " The mathematical theory of the production of radiation by collisions shows that this energy (the usual form of which is $\lambda^{-5}\phi(\lambda\theta)d\lambda$) is given by an expression of the form

$$\lambda^{-5}F(\lambda/VT)d\lambda, \qquad \ldots\ldots(8)$$

where T is the duration of a collision, V the velocity of light and F represents a function whose form depends upon the nature of the forces exerted during a collision. Comparing these two expressions we see that T must be inversely proportional to θ, that is, inversely proportional to the square of the velocity of the corpuscles. The velocity of corpuscles at 0° c. when in temperature equilibrium with their surroundings is about 10^7 cm./sec.; the wave-length at which the intensity is greatest at 0° c. is about 10^{-3} cm. In a Röntgen-ray bulb giving out hard rays, the velocity of the corpuscles may be about 10^{10} cm./sec.—or 10^3 times those in the metal; hence, if the law of duration of impacts is true, the radiation produced by the impact of the corpuscles in the tube should be a maximum for a wave-length of $10^{-3}/10^6$ or 10^{-9} cm., and [Thomson had "as"] this is of the same order as the thickness of a pulse of very penetrating Röntgen radiation; this test, as far as it goes, confirms the law of "duration of collisions". The thickness of the pulse resulting from the collision, being the distance travelled by electromagnetic radiation during the time of a collision, will also be proportional to $1/\theta$, and it is perhaps significant that if we make use of this proportionality in (7), that is, if we replace 1/L by θ, we obtain the result that ϵ is proportional to θ^4.

Thomson's conclusions were supported by the work of Jeans (1914 and 1924 a), who passed them under critical review, the conclusion being that they lead to conditions inside radiating matter most difficult to reconcile with known facts. However, this does not affect the argument which matters here, namely, that when we approach the question of the dimensions of temperature from the laws of radiation as starting point, we find that the important variable in determining changes of θ is L or T, occurring singly, rather than L/T. This result begins to suggest itself at equation (7). Thus, for radiation from a substance the specific inductive capacity of which does not vary with temperature, the only variables in (7) are L and Q. If, for example, the charge of which Q denotes the dimensions is that of a single electron, the fact that temperature influences the effective number of charges being taken care of by supposing that L includes the variable distance between electrons, then L is the sole variable. The details are unimportant for a dimensional discussion, but it may be of interest to point out that on reasonable assumptions Sir J. J. Thomson (1907 b) has included the effects of radiation between collisions, as well as at the collisions themselves, and while the latter would no doubt have been uppermost in his mind as the chief cause of radiation, he did not omit to take into account, in his second analysis of the connection between radiant energy and temperature, the part played by the path between collisions. Thus, for a dimensional discussion, L might equally well be thought of as the mean free-path, which, as Thomson points out, does in fact vary approximately as the reciprocal of the absolute temperature in the several pure metals for which the temperature coefficient of resistance is in the neighbourhood of 1/273 (Thomson 1907c).

Let us now proceed to a more rigorous dimensional discussion. Let us suppose the temperature enters as θ^4 in known manner, but let us include the possibility that charge enters as Q^a. Let A be a purely numerical constant, and let us assume for the moment that $k=1$. Then

$$\epsilon = A q^a \theta^4 \qquad \ldots\ldots.(9)$$

or, dimensionally,

$$\mathrm{M/LT^2} = \mathrm{Q}^a \Theta^4, \qquad \ldots\ldots(9\,a)$$

giving

$$\Theta^4 = \mathrm{M}^{1-\frac{a}{2}} \mathrm{L}^{-1-\frac{3a}{2}} \mathrm{T}^{a-2}. \qquad \ldots\ldots(10)$$

The only possible value of α which makes Θ independent of M is 2, for which

$$\Theta = \mathrm{L}^{-1}. \qquad \ldots\ldots(11)$$

Now this will *not* do, though it gives the correct result, as we shall see. One reason is that we have omitted to include on the left-hand side of (9 ***a***) the quantity [N]. Moreover, we have made no provision for the possibility that the mass of the exciting charge is an important factor. The justification (Thomson 1907 **d**) for the omission of M, namely, " that the expressions . . . . for the electrical and thermal conductivities, the radiation, and the other electrical effects do not involve the mass of the carrier, so that the results would hold if the carriers were bodies having a much greater mass than that of a corpuscle ", is a very strong one if we accept the classical theory of radiation. If, however, we accept the quantum theory, neither mass nor charge, but ***c***, ***h***, and ***k***, enter into the expression (O'Rahilly, 1938).

The writer has in hand a theory which seems to give radiation from uncharged as well as charged bodies. The theory has had to be shelved during the war, but for present purposes the following result is all that we require. The rate at which any mass m, whether charged or no, radiates energy is given by

$$\partial U/\partial t, \qquad \ldots\ldots(12)$$

where $-U$ is the potential energy of the mass in the field of force, and $\partial/\partial t$ denotes differentiation at a fixed point of space,—a local rate of change with the time. In case the mass m is without charge, U contains m, while if the field of force is electric, U contains m through the square of the electric charge. Thus we adhere to the Thomson-Lorentz view of electromagnetic mass, but regard mass rather than charge as the entity which is necessary for radiation. This means that in those applications for which Larmor's formula for radiation from an accelerated point-charge is reconcilable with observation, the rate of radiation may be expressed in terms of the radius a of the " point " charge, and of its mass:

$$m = 2q^2/3ac^2, \qquad \ldots\ldots(13)$$

giving

$$\partial R/\partial t = 2q^2 f^2/3c^3 = maf^2/c, \qquad \ldots\ldots(14)$$

where R is the flux of radiation and f the acceleration of the particle.

We thus make a fresh start; instead of (9) we try ($\boldsymbol{m}$ = electronic mass)

$$\epsilon = A\boldsymbol{m}^{\alpha}\theta^4. \qquad \ldots\ldots(15)$$

This time we include N, so that our dimensional equation is

$$[N]\mathrm{M}/\mathrm{LT}^2 = \mathrm{M}^{\alpha}\Theta^4. \qquad \ldots\ldots(16)$$

giving, if $\alpha = 1$,

$$\Theta^4 = [N]/\mathrm{LT}^2. \qquad \ldots\ldots(17)$$

Eliminating $[N]$ between (3) and (17) we obtain the result already reached under false assumptions at (11). The dimensions of N are thus

$$[N] = \mathrm{T}^2/\mathrm{L}^3. \qquad \ldots\ldots(18)$$

We see that by including N, and only by so doing, we have been able to reconcile (11) with the generally agreed view that temperature is energy per unit mass, generally expressed by (2), but correctly expressed by (3).

It may be argued that we had no right to take $\alpha = 1$ in reaching (17), since N is of unknown dimensions. However, N is almost certain not to contain M, and if θ is to be free of M, $\alpha = 1$ is the only alternative. If we admit the possibility that either θ or N may contain M, (16) and (3) give

$$\Theta = \mathrm{M}^{(1-\alpha)/3}/\mathrm{L}\ ; \quad [N] = \mathrm{M}^{(1-\alpha)/3}\mathrm{T}^2/\mathrm{L}^3. \qquad \ldots\ldots(19)$$

This allows those who wish to make G (the gravitational constant) dimensionless to do so, as follows. The dimensions of G are, using (19),

$$[G] = \mathrm{L}^3[N]/\mathrm{MT}^2 = \mathrm{M}^{(-2-\alpha)/3}. \qquad \ldots\ldots(20)$$

The choice is thus

$$\alpha = -2. \qquad \ldots\ldots(21)$$

This is not a choice favoured by the writer, since it involves

$$\epsilon = A\theta^4/\boldsymbol{m}^2\ ; \quad \Theta = \mathrm{M}/\mathrm{L}, \qquad \ldots\ldots(22)$$

which has no very obvious physical significance. Equation (11), however, has the immediate significance that temperature is measured by the thickness of a pulse of electromagnetic radiation (the fact that we do not make the measurement is immaterial). It should be emphasized that since the mass of the carrier is the same for all substances, assuming all the radiation due to electrons, the appearance of $\boldsymbol{m}$ in (15) is without prejudice to the experimental result that the energy density of black-body radiation is independent of the atomic weight of the substance.

It may be noticed that, according to (12),

(*a*) Acceleration is a necessary, but not a sufficient, condition for radiation.

(*b*) In particular, U would be constant for an electron moving in a circular orbit under a central force, so that despite centripetal acceleration no radiation would result. This is in accordance with observation, for the width of spectral lines is finite, which implies that the orbits are elliptical. Thus, according to the writer's theory, orbits are all circular until atoms are in some way excited, when they are distorted into ellipses. This need not necessarily conflict with the quantum theory, since there would appear to be no experiment which could establish the existence of elliptical orbits in atoms while not actually radiating. Actually, if it were not for precession of elliptical orbits (generally attributed to relativistic changes in mass) it would be possible for (12) to apply to elliptical as well as to circular orbits, for (12) states in this connection that every time the electron comes round to the same point of its orbit it finds the same field acting upon it. This condition could not be satisfied by a precessing orbit.

The above remarks are largely incidental to the discussion. We conclude this section by evaluating A in (15). If $\boldsymbol{\sigma}$ is Stefan's constant ($=5{\cdot}74\times10^{-5}$), we have

$$\epsilon=(4\boldsymbol{\sigma}/\boldsymbol{c})\theta^4. \qquad \text{......(23)}$$

Comparing with (15), taking $\alpha=1$:

$$A=4\boldsymbol{\sigma}/\boldsymbol{mc}. \qquad \text{......(24)}$$

With $\boldsymbol{m}=9{\cdot}09/10^{-28}$, $\boldsymbol{c}=3{\cdot}00\times10^{10}$,

$$A=8{\cdot}42\times10^{12}. \qquad \text{......(25)}$$

The value of A in (9) would have been $3{\cdot}32\times10^4$.

If the dimensions of temperature are the reciprocal of a length, then Stefan's and Boltzmann's constants have the dimensions of momentum and mass respectively:

$$[\boldsymbol{\sigma}]=[N]\mathrm{ML}^4/\mathrm{T}^3=\mathrm{ML/T}\ ;\ \ [\boldsymbol{k}]=[N]\mathrm{ML}^3/\mathrm{T}^2=\mathrm{M}. \qquad \text{......(26)}$$

It is hardly necessary to add that when these dimensions are inserted in the quantum formula

$$\boldsymbol{\sigma}=2\pi^5\boldsymbol{k}^4/15\boldsymbol{c}^2\boldsymbol{h}^3 \qquad \text{......(27)}$$

the dimensions of $\boldsymbol{h}$ come out as those of action, namely, $[N]\mathrm{ML}^2\mathrm{T}^{-1}=\mathrm{MT/L}$.

§ 3. THE WIEDEMANN-FRANZ RATIO AND SPECIFIC HEAT

The Wiedemann-Franz ratio is unaffected by the inclusion of N, i.e. it is still given by

$$k/\sigma=3\boldsymbol{k}^2\theta/\boldsymbol{e}^2\mathcal{J}; \qquad \text{......(28)}$$

expressing $\boldsymbol{e}$ in e.m.u. and remembering (18),

$$[k/\sigma]=\mathrm{M/L}^2; \qquad \text{......(29)}$$

the appearance of M is not of great significance, but it matters that we have L^2 in the denominator. This rather suggests that the ratio should vary as the square of the absolute temperature. The work of Meissner (see Roberts, 1928) suggests that at very low temperatures this may be true.

The thermal capacity of a body may be defined as (the number of heat units required to raise its temperature by $d\theta$) divided by $d\theta$. As usually defined, specific heat is a pure number, a ratio, and the question of its dimensions does not arise. What matters in the present connection is that thermal capacity per unit mass, to which the term specific heat is frequently, if loosely, applied, has the dimensions

$$\frac{\mathrm{H}}{\mathrm{M}\Theta}=[N]\mathrm{ML}^2\mathrm{T}^{-2}/\mathrm{M}\Theta, \qquad \ldots\ldots(30)$$

where H refers to the quantity of heat. As in connection with (28), we have taken the mechanical equivalent of heat, J, as dimensionless and have expressed H in work units. We thus have, in view of (19), which gives

$$[N]/\Theta=\mathrm{T}^2/\mathrm{L}^2, \qquad \ldots\ldots(31)$$

that thermal capacity per unit mass is, like specific heat, without dimensions. This circumstance must be regarded as fortunate in view of the widespread identification of these two quantities.

The dimensions of various quantities on the new basis are listed below:—

Temperature	1/L
Quantity of Heat	M/L
Thermal Capacity (per unit mass)	1
Thermal Capacity (per unit volume)	M/L^3
Thermal Conductivity	M/LT
Entropy	M
Enthalpy	M/L

§ 4. ADIABATIC INVARIANCE*

If F be a quantity satisfying the equation

$$F=nh, \qquad \ldots\ldots(32)$$

where n is an integer, and if a be a quantity which changes with any of the external fields of force acting on a dynamical system describing a motion conforming to the quantum condition (32), then in order that (32) may be a quantum equation in the general case when a may have any value whatever,

$$\partial F/\partial a=0. \qquad \ldots\ldots(33)$$

The function $2\overline{T}/\nu$ was shown by Boltzmann (1876) to be an adiabatic invariant (that is, to satisfy (33)).

Here $\overline{T}$ is the average kinetic energy over a period of an oscillation of frequency ν. On our present basis, all this means is that the velocity of light is an adiabatic invariant. For, regarding $\overline{T}$ as reckoned per unit mass, the dimensions of $\overline{T}$ are 1/L, the same as those of temperature, so that the dimensions of $2\overline{T}/\nu$ are those of the reciprocal of a velocity. This velocity could not possibly refer in general to the velocity of any of the mechanical parts of the system; the only possible velocity which is invariant being that of electromagnetic waves *in vacuo*, this must be the velocity concerned.

---

* See Jeans, 1924, op. cit. p. 65.

§ 5. ACKNOWLEDGEMENT

I should like to thank Professor G. Stead for his interest and valued criticism.

REFERENCES

JEANS, J., 1914. *Report on Radiation and the Quantum Theory*, p. 28 (London : The Physical Society) ; 1924 a. *Ibid.*, second edition, p. 18 ; 1924 b. *Ibid.* p. 65.
O'RAHILLY, A., 1938. *Electromagnetics*, p. 728 (London : Longmans and Co.).
ROBERTS, J. K., 1928. *Heat and Thermodynamics*, p. 238 (London : Blackie and Son).
THOMSON, J. J., 1907 a. *The Corpuscular Theory of Matter*, p. 67 (London : Arnold and Co.) ; 1907 b. *Ibid.* pp. 86 *et seq.* and 95 ; 1907 c. *Ibid.* p. 80 ; 1907 d. *Ibid.* p. 102

A PROBLEM IN TWO-DIMENSIONAL FLOW

By H. H. MACEY, M.Sc., A.R.C.S., F.Inst.P.,
The British Refractories Research Association, Stoke-on-Trent

Received 27 *October* 1941

ABSTRACT. A method of solving a particular case of the two-dimensional heat-flow equation is given and used to calculate the distribution of moisture content in a clay cylinder drying from its curved surface only. The results of such calculations are compared with experimental data.

§ 1. INTRODUCTION

It has been previously shown (Macey, 1940) how it was possible to calculate the distribution of moisture-content in clay bars drying at a constant rate from one end only. The ultimate objective of this work being the correlation of the moisture distribution with the appearance of cracks during the drying process, it was necessary to develop the theory into two or three dimensions. For practical reasons, notably the difficulties of supporting a sphere of wet clay and of ensuring uniform drying over its surface, the cylinder drying only from its curved surface is to be preferred. It was therefore required to know the moisture-content distribution at any time in such a cylinder drying at a constant rate. The problem involved is essentially similar to that of heat flow, and it may be that the method of solution adopted, though unorthodox, is capable of application in other fields.

§ 2. ONE-DIMENSIONAL FLOW

If a clay be subjected to a pressure P between porous pistons, it assumes a moisture-content (percentage on basis of dry weight of clay) M such that $P=\alpha e^{-\beta M}$, α and β being constants for any particular batch of clay. The permeability or aqueous conductivity, C, is a similar exponential function of moisture-content, $C=\theta e^{\phi M}$, θ and ϕ being again experimentally determined constants. On the assumption that the water movement is to be associated with the pressure P, the basic differential equation for linear flow is

$$\frac{\partial}{\partial x}\left[C\frac{\partial P}{\partial x}\right]=\frac{\partial}{\partial t}\left[\frac{M}{S+M}\right], \qquad \ldots\ldots(1)$$

where S is the volume of 100 grams of clay in the plastic mixture. For a detailed argument, reference must be made to the original paper, but since, experimentally, ϕ and β are approximately equal, this is reduced to

$$\alpha\beta\theta e^{(\phi-\beta)M_e}(S+M)\frac{\partial^2 M}{\partial x^2}=\frac{\partial M}{\partial t}. \qquad \ldots\ldots(2)$$

This is the familiar heat-flow equation which, by suitable adjustment of the units of time, is further reduced to

$$\frac{\partial^2 M}{\partial x^2}=\frac{\partial M}{\partial t}. \qquad \ldots\ldots(3)$$

The boundary conditions are that the moisture-content is initially uniform, that the rate of drying at the surface is constant, and that the moisture-content gradient at the non-drying end of the bar is always zero, i.e. $M=M_0$ when $t=0$, $\partial M/\partial x=A$ at the surface, and $\partial M/\partial x=0$ at the non-drying end. A solution suitable for small values of the time t is

$$M=M_0-2At^{\frac{1}{2}}\left[\chi_1\left(\frac{x}{2t}\right)+\chi_1\left(\frac{2l-x}{2t^{\frac{1}{2}}}\right)+\chi_1\left(\frac{2l+x}{2t^{\frac{1}{2}}}\right)+\chi_1\left(\frac{4l-x}{2t^{\frac{1}{2}}}\right)\right.$$
$$\left.+\chi_1\left(\frac{4l+x}{2t^{\frac{1}{2}}}\right)+\ldots\right], \qquad \ldots\ldots(4)$$

where the origin is taken at the drying surface and χ_1 is the first integral of the error-function complement as defined by Hartree (1936). The error-function as defined by Jeffreys (1927) being

$$\operatorname{erf}(\omega)=\frac{2}{\sqrt{\pi}}\int_0^{\omega}e^{-\xi^2}\,d\xi, \qquad \ldots\ldots(5)$$

Hartree takes the complement to be the function

$$\chi(\omega)=\frac{2}{\sqrt{\pi}}\int_{\omega}^{\infty}e^{-\xi^2}\,d\xi, \qquad \ldots\ldots(6)$$

and repeated integrals are defined as

$$\chi_n(\omega)=\int_{\omega}^{\infty}\chi_{n-1}(\xi)\,d\xi, \qquad \ldots\ldots(7)$$

with $\chi_0(\omega)=\chi(\omega)$ as in (6).

A second solution suitable for large times is

$$M=M_0-\frac{A}{l}\left[\left(t+\frac{x^2}{2}-\frac{l^2}{6}\right)-\frac{2l^2}{\pi^2}\sum_{n=1}^{\infty}(-1)^n\frac{1}{n^2}\cos\frac{n\pi x}{l}\,.\,e^{-n^2\pi^2t/l^2}\right], \qquad \ldots\ldots(8)$$

in which the origin is taken at the non-drying end of the bar. For large values of t, this reduces to

$$M=M_0-\frac{A}{l}\left(t+\frac{x^2}{2}-\frac{l^2}{6}\right). \qquad \ldots\ldots(9)$$

These forms of solution are highly convenient, for when more than two terms of (4) are necessary, the simple solution (9) holds, and it is never required to evaluate the summation term in (8).

§ 3. TWO-DIMENSIONAL FLOW

With the same basic theory, and taking the origin at the centre, the differential equation in the case of a cylinder may be written

$$\alpha\beta\theta\, e^{(\phi-\beta)M_c}(S+M)\frac{\partial}{dx}\left[x\frac{\partial M}{dx}\right]=x\frac{\partial M}{\partial t}, \qquad \ldots\ldots(10)$$

which, by adjustment of the units of time as before, reduces to

$$\frac{\partial}{\partial x}\left[x\frac{\partial M}{\partial x}\right]=x\frac{\partial M}{\partial t}. \qquad \ldots\ldots(11)$$

A strict method of solving this, under the same boundary conditions, is unknown.

A further condition to be observed is that the total change of moisture-content at any time must be equal to the loss of water by drying, i.e.,

$$\pi M_0 l^2-2\pi\int_0^l Mx\,dx=2\pi Al\tau, \qquad \ldots\ldots(12)$$

where τ is the time of drying of the cylinder of radius l.

A particular solution, valid for large times, and satisfying all these conditions, is

$$M=M_0-\frac{A}{l}\left[2\tau+\frac{x^2}{2}-\frac{l^2}{4}\right]. \qquad \ldots\ldots(13)$$

Compare this with the solution (9) for a bar of length l. It is at once apparent that they are identical if

$$2\tau-\frac{l^2}{4}=t-\frac{l^2}{6}.$$

In both the bar of length l and the cylinder of radius l, drying at the same rate, the moisture-content is initially uniform at M_0. At large times, the curves of moisture-content distribution are identical parabolas *on a different time scale*, which do not change in shape, but only in position along the moisture-content scale with time. The solution for small times now rests upon the following argument: if two families of curves of similar derivation, both subject to the same boundary conditions which define the constant values of their slopes at their limits, are initially straight lines and are ultimately identical parabolas, then they must have the same evolution. The solution (4) for linear flow must also hold for the cylinder, but at some time other than t. Knowing the time of drying τ of a cylinder of radius l, the problem reduces to that of finding the equivalent time of drying t of a bar of length l, when that bar will have the same moisture-content distribution as the cylinder.

This can be done by using the condition (12) and, transferring the origin to the surface of the cylinder, we have

$$2\pi l A\tau = 4\pi\int_0^l At^{\frac{1}{2}}\left[\chi_1\left(\frac{x}{2t^{\frac{1}{2}}}\right)+\chi_1\left(\frac{2l-x}{2t^{\frac{1}{2}}}\right)+\chi_1\left(\frac{2l+x}{2t^{\frac{1}{2}}}\right)+\dots\right](l-x)\,dx. \qquad \dots\dots(14)$$

The integral is derived as follows:

We have, integrating by parts,

$$\int_\omega^\infty \xi\chi_{n-1}(\xi)d\xi = \omega\chi_n(\omega)+\chi_{n+1}(\omega). \qquad \dots\dots(15)$$

and in particular

$$\int_\omega^\infty \xi\chi_1(\xi)d\xi = \omega\chi_2(\omega)+\chi_3(\omega).$$

Hence if ξ is a function of x

$$\int_{x=0}^{x=l} \xi\chi_1(\xi)d\xi = \Big[\xi\chi_2(\xi)+\chi_3(\xi)\Big]_{x=l}^{x=0}. \qquad \dots\dots(16)$$

It is required to evaluate

$$2t^{\frac{1}{2}}\int_0^l\left[\chi_1\left(\frac{x}{2t^{\frac{1}{2}}}\right)+\chi_1\left(\frac{2l-x}{2t^{\frac{1}{2}}}\right)+\chi_1\left(\frac{2l+x}{2t}\right)+\chi_1\left(\frac{4l-x}{2t^{\frac{1}{2}}}\right)+\ \dots\right](l-x)dx. \qquad \dots\dots(17)$$

Now

$$\begin{aligned}2t^{\frac{1}{2}}l\int_0^l\left[\chi_1\left(\frac{x}{2t}\right)+\chi_1\left(\frac{2l-x}{2t}\right)+\ \dots\right]dx \\ &= 4tl\left[\chi_2\left(\frac{x}{2t^{\frac{1}{2}}}\right)-\chi_2\left(\frac{2l-x}{2t^{\frac{1}{2}}}\right)+\chi_2\left(\frac{2l+x}{2t^{\frac{1}{2}}}\right)+\ \dots\right]_l^0 \\ &= 4tl\ \chi_2(0) \\ &= tl \qquad \dots\dots(18)\end{aligned}$$

since $\chi_2(0)=\frac{1}{4}$.

To evaluate the second term of (17) put $\xi = x/2t^{\frac{1}{2}}$ in (16).

Then

$$\int_0^l\left(\frac{x}{2t^{\frac{1}{2}}}\right)\chi_1\left(\frac{x}{2t^{\frac{1}{2}}}\right)\left(\frac{dx}{2t^{\frac{1}{2}}}\right)=\left[\left(\frac{x}{2t^{\frac{1}{2}}}\right)\chi_2\left(\frac{x}{2t^{\frac{1}{2}}}\right)+\chi_3\left(\frac{x}{2t^{\frac{1}{2}}}\right)\right]_l^0$$

or

$$\int_0^l x\chi_1\left(\frac{x}{2t^{\frac{1}{2}}}\right)dx = 4t\left[\chi_3(0)-\left(\frac{l}{2t}\right)\chi_2\left(\frac{l}{2t^{\frac{1}{2}}}\right)-\chi_3\left(\frac{l}{2t^{\frac{1}{2}}}\right)\right]. \qquad \dots\dots(19)$$

Similarly, putting $\xi = \left(\frac{2l-x}{2t^{\frac{1}{2}}}\right)$ in (16), we find

$$\begin{aligned}\int_0^l\left(\frac{2l-x}{2t^{\frac{1}{2}}}\right)\chi_1\left(\frac{2l-x}{2t^{\frac{1}{2}}}\right)\left(\frac{dx}{2t^{\frac{1}{2}}}\right) &= \left(\frac{l}{2t^{\frac{1}{2}}}\right)\chi_2\left(\frac{l}{2t^{\frac{1}{2}}}\right) \\ &+\chi_3\left(\frac{l}{2t^{\frac{1}{2}}}\right)-\left(\frac{2l}{2t^{\frac{1}{2}}}\right)\chi_2\left(\frac{2l}{2t^{\frac{1}{2}}}\right)-\chi_3\left(\frac{2l}{2t^{\frac{1}{2}}}\right). \qquad \dots\dots(20)\end{aligned}$$

But

$$\int_0^l \left(\frac{2l-x}{2t^{\frac{1}{2}}}\right) \chi_1\left(\frac{2l-x}{2t^{\frac{1}{2}}}\right) \frac{dx}{2t_{\frac{1}{2}}} = \frac{1}{4t}\int_0^l 2l\chi_1\left(\frac{2l-x}{2t^{\frac{1}{2}}}\right) dx - \frac{1}{4t}\int_0^l x\chi_1\left(\frac{2l-x}{2t^{\frac{1}{2}}}\right) dx$$

and

$$\frac{1}{4t}\int_0^l 2l\chi_1\left(\frac{2l-x}{2t^{\frac{1}{2}}}\right) dx = -\left(\frac{2l}{2t^{\frac{1}{2}}}\right)\left[\chi_2\left(\frac{2l-x}{2t^{\frac{1}{2}}}\right)\right]_l^0 = \left(\frac{2l}{2t^{\frac{1}{2}}}\right)\chi_2\left(\frac{l}{2t^{\frac{1}{2}}}\right) - \left(\frac{2l}{2t^{\frac{1}{2}}}\right)\chi_2\left(\frac{2l}{2t^{\frac{1}{2}}}\right). \qquad \ldots\ldots(21)$$

Therefore

$$\frac{1}{4t}\int_0^l x\chi_1\left(\frac{2l-x}{2t^{\frac{1}{2}}}\right) dx = \left(\frac{l}{2t^{\frac{1}{2}}}\right)\chi_2\left(\frac{l}{2t^{\frac{1}{2}}}\right) + \chi_3\left(\frac{2l}{2t^{\frac{1}{2}}}\right) - \chi_3\left(\frac{l}{2t^{\frac{1}{2}}}\right). \qquad \ldots\ldots(22)$$

Similarly, put $\xi = \left(\frac{2l+x}{2t^{\frac{1}{2}}}\right)$. Then

$$\frac{1}{4t}\int_0^l x\chi_1\left(\frac{2l+x}{2t^{\frac{1}{2}}}\right) dx = \chi_3\left(\frac{2l}{2t^{\frac{1}{2}}}\right) - \left(\frac{l}{2t^{\frac{1}{2}}}\right)\chi_2\left(\frac{3l}{2t^{\frac{1}{2}}}\right) - \chi_3\left(\frac{3l}{2t^{\frac{1}{2}}}\right). \qquad \ldots\ldots(23)$$

In the same manner

$$\frac{1}{4t}\int_0^l x\chi_1\left(\frac{4l-x}{2t^{\frac{1}{2}}}\right) dx = \chi_3\left(\frac{4l}{2t^{\frac{1}{2}}}\right) + \left(\frac{l}{2t^{\frac{1}{2}}}\right)\chi_2\left(\frac{3l}{2t^{\frac{1}{2}}}\right) - \chi_3\left(\frac{3l}{2t^{\frac{1}{2}}}\right), \qquad \ldots\ldots(24)$$

and similarly for subsequent terms.

Adding (19), (22), (23), (24),

$$\int_0^l \left[\chi_1\left(\frac{x}{2t^{\frac{1}{2}}}\right) + \chi_1\left(\frac{2l-x}{2t^{\frac{1}{2}}}\right) + \chi_1\left(\frac{2l+x}{2t^{\frac{1}{2}}}\right) + \ldots\right] x\,dx$$

$$= 4t\left[\chi_3(0) - 2\chi_3\left(\frac{l}{2t^{\frac{1}{2}}}\right) + 2\chi_3\left(\frac{2l}{2t^{\frac{1}{2}}}\right) - 2\chi_3\left(\frac{3l}{2t^{\frac{1}{2}}}\right) + \ldots\right]$$

$$= 4t\left[\chi_3(0) + 2\Sigma(-1)^n\chi_3\left(\frac{nl}{2t^{\frac{1}{2}}}\right)\right]. \qquad \ldots\ldots(25)$$

Hence, from (14), (17), (18) and (25),

$$l\tau = lt - 8t^{3/2}\left[\chi_3(0) + 2\Sigma(-1)^n\chi_3\left(\frac{nl}{2t^{\frac{1}{2}}}\right)\right]. \qquad \ldots\ldots(26)$$

Thus, knowing τ, (26) may be solved for t, which value is used in (4) to give the moisture-content distribution.

In general the summation term in (26) is small compared with $\chi_3(0) = 0{\cdot}094$, increasing relatively for the larger times. In the particular cases encountered it did not exceed 4 % and could be neglected, whereupon (26) reduces to a cubic and can be solved in the usual manner. It was found more convenient, however, to construct a table giving τ in terms of t, and to obtain particular values by interpolation.

§ 4. EXPERIMENTAL CONFIRMATION

Confirmatory experimental evidence, similar to that described in the paper to which reference has already been made, was obtained by determining actual moisture-content distributions both in bars of uniform cross-section and in

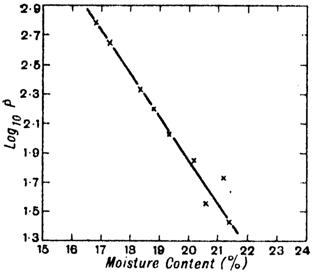


Figure 1.

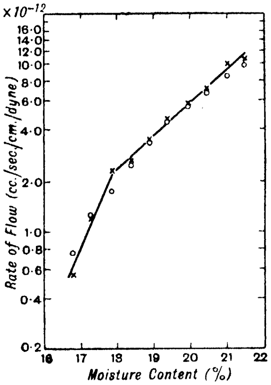


Figure 2.

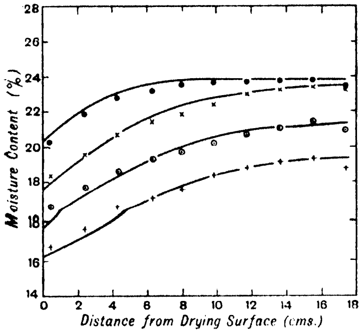


Figure 3.

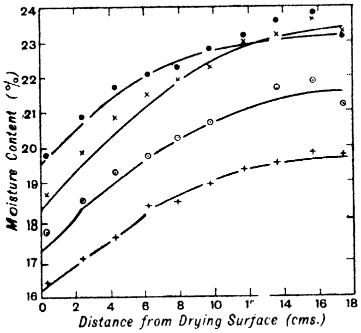


Figure 4.

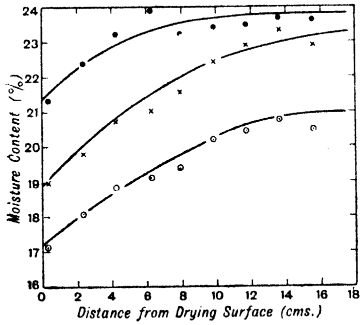


Figure 5.

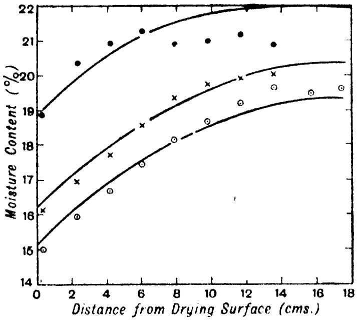


Figure 6.

Figure 1. Pressure-moisture content relation of clay B.
Figure 2. Conductivity of clay B.
Figures 3, 4. Comparison of experimental and calculated moisture-content distributions in the case of linear flow.
Figures 5, 6. Comparison of experimental and calculated moisture-content distributions in the case of radial flow.

wedge-shaped pieces, suitably covered to prevent drying except from their exposed ends. The experimental method of drying the testpieces under controlled conditions of temperature and humidity, and of cutting sections along their length after a noted time, calls for no particular comment. The clay used was a finely ground fireclay, known as Clay B, the relation between pressure and equilibrium moisture-content being given in figure 1, and the aqueous conductivity in figure 2. The characteristic constants of the clay are

$\alpha = 3{\cdot}422 \times 10^{12}$ $\quad S = 39{\cdot}8$

$\beta = 0{\cdot}6733$ $\quad M_L = 11{\cdot}5\%$

θ (above $17{\cdot}8\ \%$) $= 9{\cdot}116 \times 10^{-16}$ $\quad V_D = 51{\cdot}0$ c.c.

ϕ (above $17{\cdot}8\ \%$) $= 0{\cdot}4376$

In the above, M_L is the moisture-content at which shrinkage ceases (and below which the above theory is not valid) and V_D is the dry volume of 100 grams of the clay. From these and S, the volumetric shrinkage curve may be reconstructed.

Typical moisture-content distributions in bars are shown in figures 3 and 4, the rates of drying of those in figure 3 being, from top to bottom, 9·4, 8·5, 6·9 and $7{\cdot}25 \times 10^{-6}$ gm. sec.$^{-1}$ cm.$^{-2}$ respectively. The corresponding rates in figure 4 were 6·0, 6·5, 5·8, and $5{\cdot}7 \times 10^{-6}$ respectively. The points shown are the experimental values, and the curves were calculated by the method described above.

Moisture-content distributions in wedges are shown in figures 5 and 6, the rates of drying being: for figure 5, 5·1, 5·5 and $6{\cdot}7 \times 10^{-6}$ and for figure 6, 7·8, 10·1 and $10{\cdot}8 \times 10^{-6}$ gm. sec.$^{-1}$ cm.$^{-2}$ respectively. On cutting thin slices from the thinner end of the wedge, very small quantities of clay are available for moisture-content determination, and the experimental accuracy is naturally less in this case.

§ 5. THREE-DIMENSIONAL FLOW

It may be noted that a parabolic expression similar to (9) and (13) is a solution in the case of three-dimensional flow. On the same argument, (4) is also a solution on another different time scale.

§ 6. ACKNOWLEDGEMENTS

The author wishes to express his thanks to the Council and Director of the British Refractories Research Association for their permission to publish this paper, and to Mr. F. G. Wilde, who carried out most of the experimental work involved.

REFERENCES

HARTREE, D. R., 1936. *Mem. Manchr. Lit. Phil. Soc.*

JEFFREYS, H., 1927. *Operational Methods in Mathematical Physics* (Cambridge: University Press).

MACEY, H. H., 1940. *Proc. Phys. Soc.* **52**, 625.

ON SPACE-CHARGE EFFECTS IN VELOCITY-MODULATED ELECTRON BEAMS

By W. H. J. FUCHS and R. KOMPFNER

Communicated by W. E. Benham, 25 *September* 1941

ABSTRACT. For the purposes of analysis, a velocity-modulated beam of electrons is idealized in two different ways. The cross-section of the beam is assumed very large. In the first model the electrons move in a drift space enclosed by parallel planes having equal potential and screening the beam from the rest of space. The acceleration of an electron in the drift space is entirely due to space-charge forces. This mathematical model is used to discuss the working of velocity-modulated tubes as high-frequency amplifiers. The main conclusion is that the beam of electrons operates in the same manner as if space charge were absent, except that the mean transit-time is altered. The saturation current-density obeys a $V^{3/2}$ law in its dependence on the accelerating voltage V. For given j_0 and V, the length of the tube must be below a critical value in order that a continuous beam of electrons can exist. The limiting length of an electron beam (subject to the above assumptions) is just under three times the plate separation of a saturated plane diode passing the same current per cm^2.

The analysis further shows the existence of a potential hill with its crest about half-way across the drift space. Electrons entering the space are slowed down by the space charge in front of them until they reach the summit of the potential hill; from there on they are accelerated again by the space charge behind them.

The second model beam, called the *domain* model, if anything under-estimates the effects of space charge. A domain may be thought of as a thin slice of beam which moves under the influence of its own initial distribution of charge and velocity, the effects of neighbouring domains and of the boundaries being neglected. In the course of the motion of a domain, two kinds of charge-density maxima will occur. The first kind (" bunching ") occurs when and where successive layers of electrons overtake one another. The second kind (" compression ") occurs while no overtaking takes place. Successive layers of electrons approach each other to a minimum distance and then move further apart again. The conditions for bunching are found.

---

§ 1. INTRODUCTION

THE purpose of this paper is the discussion of two mathematical models of velocity-modulated beams of electrons. The study of such beams has recently attracted considerable attention. D. L. Webster (1939) has given a complete mathematical analysis under the assumptions that the modulation is small and sinusoidal and that the influence of space charge may be neglected. His paper also contains a calculation concerning the effect of space charge. This calculation is based on several assumptions, some of them rather remote from physical reality. In particular it is suggested that Webster's assumption that the average, or unmodulated, value of space-charge density is zero (as if cancelled by positive ions) prevents a proper estimate of the importance of space charge being made. It may be of interest to discuss the effect of space charge under different assumptions. We have constructed two different mathematical models of velocity-modulated beams.

We apply the first model, which is discussed in § 2, to the study of drift tubes working as amplifiers at very high frequencies. The second model, described in § 3, is devised so as to allow a very simple analysis of the process of "bunching" in the presence of space charge.

Throughout the paper we make the following assumptions:—

(i) The beam is of infinite cross-section and all electrons move along parallel lines, so that the problem becomes one-dimensional.

(ii) Electro-magnetic and relativistic effects can be neglected. This assumption is justified except for very fast electrons.

§ 2. DISCUSSION OF THE FIRST MODEL

2.1. *Assumptions and notation.* In this section we make the following additional assumptions:—

(*a*) We consider the motion of electrons between two infinite planes $x=0$ and $x=l$. The electrons enter the space between the planes through $x=0$ with the velocity $v_0(t)$ and fly across it to $x=l$ along straight lines parallel to the x-axis. The drift-space between $x=0$ and $x=l$ is electrostatically screened against the rest of space.

(*b*) The current-density at $x=0$ is constant and equal to $-j_0$ in the direction of increasing x.

(*c*) There is no overtaking of electrons.

(*d*) $v_0(t)=v_0+v_1(t)$, where v_0 is constant and $v_1(t)$ is a periodic modulation. The mean value of $v_1(t)$ over a period is 0.

(*e*) The mean transit time of electrons through the space we call T. We write $h(t)=T(1+\eta(t))$ for the time of transit of an electron which leaves the space at the time t. Because of (*d*), $\eta(t)$ is periodic, and by the definition of T the mean value of $\eta(t)$ over a period is 0. From 2.3 on, we shall assume that $\eta(t)$ is small.

For convenience we give the following list of our symbols:—

Time: t.

Charge of electron: $-e$. Mass of electron: m.

Charge-density: $-\rho(x, t)$.

Velocity of electron which entered the space at $t=t_1$: $v(t, t_1)$.

Distance traversed by this electron: $x(t, t_1)$.

Time of transit through the space of the electron passing $x=l$ at the moment t : $h(t)$.

Current-density at $x=0$: $-j_0$.

Velocity of electron at $x=0$: $v_0(t)=v_0+v_1(t)$.

$$\beta=2\pi(\boldsymbol{e}/\boldsymbol{m})j_0.$$

Electrostatic and c.g.s. units are employed throughout the calculation, unless stated otherwise.

2.2. *The equation of motion.* The force exerted on an electron by the space charge between the planes x and $x+dx$ is the same as that due to a condenser plate carrying a charge $-\rho(x, t)dx$ per unit area. This force is $2\pi e\rho(x, t)dx$, repelling the electron from the plate. Hence the total force on the electron at $x=u$ which entered the space at $t=t_1$, say, is

$$\boldsymbol{m}\frac{dv(t,t_1)}{dt}=2\pi\boldsymbol{e}\left\{\int_0^u \rho(x,t)dx-\int_u^l \rho(x,t)dx\right\}. \qquad \ldots\ldots(1)$$

This equation can be written in a different form, if we introduce the transit-time $h(t)$ of an electron. The first integral in (1) is, apart from sign, the charge per unit cross-section to the left of the electron (with the usual orientation of the x-axis). That is to say, it is the charge which entered the space between the time t_1 and the time t. But this charge is equal to $-j_0(t-t_1)$ by assumption (b). Hence

$$\int_0^u \rho(x,t)dx=j_0(t-t_1).$$

Similarly

$$\int_u^l \rho(x,t)dx=j_0(t_1-t+h(t)) \text{ (see figure 1).}$$

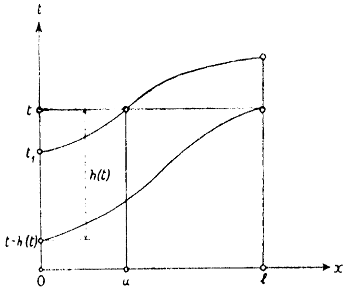


Figure 1. Space-time diagram of electron motion.

Therefore we can re-write (1) as

$$\frac{dv(t,t_1)}{dt}=\beta\{2(t-t_1)-h(t)\}, \qquad \ldots\ldots(1a)$$

where $\beta=2\pi(\boldsymbol{e}/\boldsymbol{m})j_0$. Two successive integrations give

$$x(t,t_1)=\beta\left\{\tfrac{1}{3}(t-t_1)^3-\int_{t_1}^t du\int_{t_1}^u h(v)dv\right\}+v_0(t_1)(t-t_1). \qquad \ldots\ldots(2)$$

If $x(t, t_1) = l$, we can write $t - h(t)$ instead of t_1 by the definition of $h(t)$. After a simple transformation (2) now becomes

$$l = \beta\left\{\tfrac{1}{3}h^3(t) - \int_0^{h(t)} zh(t-z)dz\right\} + v_0(t - h(t)) \,.\, h(t). \qquad \ldots\ldots(3)$$

If $h(t)$ has been found from (3), the motion is completely determined. If, for instance, we want to find the current-density at $x = l$, we need only remember that the charge passing through $x = l$ per unit area in the time interval between t and $t + \delta t$ is the charge which entered the space between the time $t - h(t)$ and $t + \delta t - h(t + \delta t)$, so that the current-density at $x = l$ is

$$\lim_{\delta t \to 0}\left[-j_0\frac{\delta t - h(t+\delta t) + h(t)}{\delta t}\right] = -j_0(1 - h'(t)). \qquad \ldots\ldots(4)$$

2.3. *Simplification of the equation of motion for small* $\eta(t)$

We assume from now on that the variable part $T\eta(t)$ of $h(t)$ is small compared with the mean transit time T. This will be the case if, and only if, the modulation $v_1(t)$ is small compared with v_0. We shall from now on neglect terms of the order of $\eta^2(t)$ and the product $v_1(t)\eta(t)$ in comparison with terms of the order of unity. Writing $h(t) = T[1 + \eta(t)]$, and omitting small terms, we obtain from (3)

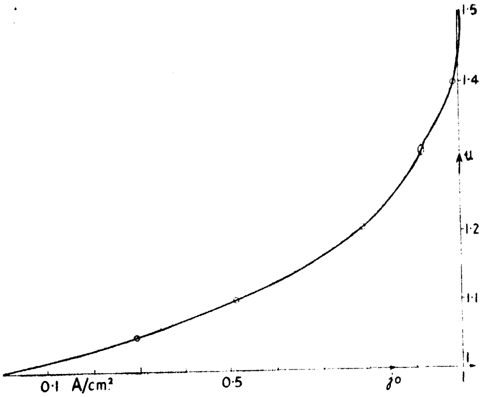


Figure 2. $j_0 - u$.

j_0	0·287	0·51	0·79	0·93	0·994	1·004
u	1·05	1·1	1·2	1·3	1·4	1·5

$$\begin{aligned} l &= \beta\{\tfrac{1}{3}T^3 + T^3\eta(t) - \tfrac{1}{2}T^3 - T^3\eta(t) - T\int_0^T z\eta(t-z)dz\} \\ &\quad + v_0T + v_0T\eta(t) + Tv_1[t - h(t)] \\ &= \beta\{-\tfrac{1}{6}T^3 - T\int_0^T z\eta(t-z)dz\} + v_0T + v_0T\eta(t) + Tv_1\{t - h(t)\}. \end{aligned}$$

$$\ldots\ldots(5)$$

We split up (5) into the two equations,

$$l = -\tfrac{1}{6}\beta T^3 + v_0 T \qquad \ldots\ldots(6)$$

and

$$0 = v_0 T\eta(t) + Tv_1[t-h(t)] - \beta T\int_0^T z\eta(t-z)\,dz. \qquad \ldots\ldots(7)$$

(6) is an equation for the mean transit time T. It shows that the presence of space charge increases T above its value $T' = l/v_0$, which holds when there is space charge.

For the ratio $T/T' = u$ we have the equation

$$-\tfrac{1}{3}au^3 + u = 1, \qquad \ldots\ldots(8)$$

where a is the dimensionless number $\frac{1}{2}\beta l^2 v_0^{-3}$. The left side of (8) has a maximum at $u = a^{-\frac{1}{2}}$ where it is equal to $\frac{2}{3}a^{-\frac{1}{2}}$. Therefore (8) has two positive roots as long as $\frac{2}{3}a^{-\frac{1}{2}} > 1$, i.e. $a < 4/9$. Only the smaller of these roots has physical significance. As a increases from 0 to its largest possible value 4/9, the influence of space charge becomes more and more appreciable until u is increased from its " no-space-charge " value 1 to 1·5. If a is increased beyond 4/9, the equation (8) no longer has positive roots and an electron beam cannot be formed. We have, therefore, the following condition for the existence of a continuous beam:—

$$4/9 \geq a = \pi(\boldsymbol{e}/\boldsymbol{m})j_0 l^2 v_0^{-3} \quad \text{or} \quad l \leq \tfrac{2}{3}(2v_0^3/\beta)^{\frac{1}{2}}, \quad \text{or} \quad j_0 \leq \boldsymbol{m}v_0^3 l^{-2}/9\pi\boldsymbol{e}.$$

Again,

$$T' < T \leq \tfrac{3}{2}T' \leq \left(\frac{2v_0}{\beta}\right)^{\frac{1}{2}}. \qquad \ldots\ldots(9)$$

If v_0 is expressed in terms of the accelerating voltage $V_0(v_0 = (2\boldsymbol{e}V_0/\boldsymbol{m})^{\frac{1}{2}})$, the critical length and current are given by

$$l = \left\{\frac{8\sqrt{(2\boldsymbol{e}/\boldsymbol{m})}}{9\pi j_0} V_0^{3/2}\right\}^{\frac{1}{2}}, \quad j_0 = \frac{8\sqrt{(2\boldsymbol{e}/\boldsymbol{m})}}{9\pi l^2} V_0^{3/2}, \qquad \ldots\ldots(9')$$

so that we have a $V^{3/2}$ law. It is interesting to compare (9′) with Langmuir's formula,

$$L = \left\{\frac{\sqrt{(2\boldsymbol{e}/\boldsymbol{m})}}{9\pi I} V^{3/2}\right\}^{\frac{1}{2}},$$

connecting the plate distance L, the saturation current I per cm.^{2} and the applied voltage V of a plane diode. If $V = V_0$,

$$l/L = 2\sqrt{(2I/j_0)}.$$

In particular, the critical length of an electron beam is $2\sqrt{2}$ times the plate separation of a Langmuir diode passing the same saturation current at the same accelerating voltage.

It may be instructive to examine in more detail the manner in which the space charge increases the transit time. We confine ourselves to the case of a steady, unmodulated current. This special case shows all the relevant features.

For a steady current $v_0(t)=v_0$ and $h(t)=T$. Using these values in (1*a*) with $t_1=0$ (no loss of generality) and integrating twice we obtain

$$v(t)=v(t,0)=\beta\{t^2-Tt\}+v_0,$$
$$x(t)=x(t,0)=\beta\{\tfrac{1}{3}t^3-\tfrac{1}{2}Tt^2\}+v_0t.$$

The potential $-V$ at any point is given by the equation of energy,

$$\tfrac{1}{2}\boldsymbol{m}\boldsymbol{v}^2+\boldsymbol{e}\boldsymbol{V}=\tfrac{1}{2}\boldsymbol{m}\boldsymbol{v}_0^2.$$

The function $v(t)$ decreases from v_0 to a minimum at $t=\frac{1}{2}T$, i.e. half-way across the drift space. From there on, the velocity increases again until it reaches its original value v_0 at $x=l$. Correspondingly there is a " potential hill " with its crest at the middle of the drift space. Electrons are slowed down while going up the hill and gain speed while going down on the other side. In the critical case, the reduction of speed at the crest of the hill is as much as 50 %.

Numerical example:—Suppose $l=2$ cm., $v_0=3600$ volts. Then the critical current-density $j_0=1{\cdot}004$ amp./cm., minimum velocity = 900 v. The increase of u with j_0 is shown in figure 2.

2.4. *Discussion of a special case.* In general it will not be possible to find $\eta(t)$ explicitly from (7). We restrict ourselves further in discussion to the special case where

$$v_1(t)=\tfrac{1}{2}Mv_0\sin\omega t.$$

By our previous assumptions, M must be small. This will certainly be the case for a drift tube working as an amplifier, where we may reasonably suppose that M is of the order of 10^{-2}. We are mainly interested in oscillations of very high frequency, for which we may assume that the mean transit time T is not much below the time of two complete cycles. The period of a cycle is $2\pi/\omega$, so that we shall have $T(2\pi/\omega)^{-1}$ not much below 2, or, say,

$$\omega T\geqslant 10.$$

We may assume that j_0 is of the order of 0·1 ma. This value is probably below those employed in practice, but, on the other hand, our assumption of an infinite cross-section over-estimates the influence of space charge very considerably. With $j_0=0{\cdot}1$ ma. we should have

$$\beta=9.\,10^{23}\text{ cm./sec.}^{-3},$$

approximately. For v_0 and ω it will be reasonable to assume 2.10^9 and 10^9 respectively.

We assume for the moment that it will be possible to find a first approximation to $\eta(t)$ by neglecting the integral in (7) altogether. We shall see afterwards whether this assumption is justified. (7) is now reduced to

$$\eta(t)=-\tfrac{1}{2}M\sin\omega(t-T-T\eta(t)). \qquad \ldots\ldots(7a)$$

In this equation β no longer occurs explicitly, and it is easy to see that (7*a*) is valid in the case of no space charge, if we take for T the value $T'=l/v_0$. We can find $\eta(t)$ explicitly by applying the following special case of

Lagrange's theorem. If $y = r\phi(y)$, $\phi(0) = 0$, then

$$y = \sum_{n=1}^{\infty} \frac{r^n}{n!} \left(\frac{d}{dy}\right)^{n-1} [\phi(y)]^n_{y=0}.$$

With

$$y = \omega T\eta(t), \quad x = \omega(t - T), \quad r = -\tfrac{1}{2}\omega TM, \quad \phi(y) = \sin(x - y)$$

we obtain, after some re-arrangements,

$$\eta(t) = -\frac{2}{\omega T} \sum_{m=1}^{\infty} \frac{1}{m} \mathcal{J}_m(mr) \sin mx, \qquad \ldots\ldots(10)$$

where $\mathcal{J}_m(x)$ is the Bessel function of order m.*

We must now investigate whether this solution represents a sufficiently good approximation. With the numerical values of β and v_0 given above we have, by (9),

$$T < 6{\cdot}7 \,.\, 10^{-8}\,\text{sec.} \qquad \ldots\ldots(11)$$

Hence $|r| = \frac{1}{2}M\omega T < 0{\cdot}5$, say. For $|r| < \frac{1}{2}$, a fair approximation to the right-hand side of (10) is

$$\eta(t) = -\tfrac{1}{2}M \sin x. \qquad \ldots\ldots(12)$$

If we use this expression for $\eta(t)$ we find

$$\int_0^T z\eta(t-z)dz = -\tfrac{1}{2}(MT/\omega)\cos(x - \omega T) - \tfrac{1}{2}M\omega^{-2}(\sin(x - \omega T) - \sin x),$$

so that the integral in (7) is indeed small compared with the first two terms, since

$$\beta T/\omega v_0 < 3 \,.\, 10^{-2}, \quad \beta/v_0\omega^2 < 5 \,.\, 10^{-4}.$$

The approximation (12) is, therefore, quite satisfactory, as its deviation from the true value of $\eta(t)$ will be well below 10 %.

By (4) and (12), the current density at $x = l$ is given by

$$j(l) = j_0(1 + \tfrac{1}{2}M\omega T \cos\omega(t - T)). \qquad \ldots\ldots(13)$$

2.5. *Conclusion.* If we suppose the velocity of the electrons due to an applied voltage U modulated by a signal $MU \sin\omega t$, then (13) shows that the beam-current magnification-factor is $\frac{1}{2}\omega T$. This formula is also true in the case of no space charge, if $T = T' = l/v_0$. In this case the factor can be made arbitrarily large by choosing l large. If the space charge is taken into consideration, the magnification factor is limited by (9). With our numerical values, the maximum is about 30, which would be attained in a drift tube of about 80 cm. We do not want to lay stress on this numerical result, however.

The main point of our calculation is that *below a certain critical length, velocity-modulated beams of electrons can operate in the same manner as beams of velocity-modulated particles on which no space charge is acting. The chief effect of space charge is an increase in the mean transit time. Formulae valid in the case*

---

* This result also follows from Webster's work, loc. cit.

of no space charge represent good first approximations if $T' = l/v_0$ is replaced by the actual mean transit time T. Above the critical length, the space charge prevents the formation of a continuous beam.

When the current density is high, β is very large and the maximum attainable value of T as calculated from (9) becomes very small. On the other hand, the beam-current magnification is proportional to T, and it is therefore desirable to give T a large value. This cannot be done by a mere lengthening of the tube, because of the space charge. It seems, therefore, advantageous to reduce the effect of space charge. The easiest way to do this would be to leave some gas in the drift tube, which would yield positive ions under electron bombardment.

In the opinion of the authors, the qualitative conclusions reached are also valid for beams of finite cross-section. This opinion is supported by the following calculation, which establishes the existence of a " potential hill " for an (unmodulated) electron beam of cylindrical cross-section. In such a beam we shall therefore have a slowing up of the electrons and a consequent increase of the mean transit-time T.

Suppose this were not the case. Then electrons would move parallel to the axis of the beam at a nearly constant velocity v_0 and the beam would represent a cylinder of space charge of circular cross-section, radius a, say, and of length l (we neglect the gradual broadening of the beam due to space charge). The density of the space charge would be $-\rho = -j_0/v_0$, uniform throughout the cylinder. Let V be the potential at $x=0$, V_1 the potential at $x=\frac{1}{2}l$, half-way along the beam, both measured on the axis of the beam. Let v_1 be the velocity of an electron at $x=\frac{1}{2}l$. By our assumption, v_0 and v_1 are practically equal. On the other hand, an elementary integration shows that

$$\begin{aligned}\tfrac{1}{2}\boldsymbol{m}(v_0{}^2 - v_1{}^2) &= -\boldsymbol{e}(V_1 - V_0)\\ &= 2\pi\rho\boldsymbol{e}\int_{-k}^{k} du \int_0^a \left[\frac{r\,dr}{\{r^2+u^2\}^{\frac{1}{2}}} - \frac{r\,dr}{\{r^2+(k-u)^2\}^{\frac{1}{2}}}\right]\\ &= 2\pi\rho\boldsymbol{e}\{k^2 + k(a^2+k^2)^{\frac{1}{2}} - k(a^2+4k^2)^{\frac{1}{2}}\\ &\qquad + \tfrac{1}{2}a^2\log[k+(a^2+k^2)^{\frac{1}{2}}] + \tfrac{1}{2}a^2\log[-2k+(a^2+4k^2)_{\frac{1}{2}}]\\ &\qquad - \tfrac{1}{2}a^2\log[-k+(a^2+k^2)^{\frac{1}{2}}] - \tfrac{1}{2}a^2\log a\},\end{aligned}$$

where $k=\frac{1}{2}l$. If a is small compared with k, this is approximately equal to

$$\left.\begin{aligned}\tfrac{1}{2}\boldsymbol{m}(v_0{}^2 - v_1{}^2) &= \pi\boldsymbol{e}\rho a^2(\tfrac{1}{2} + \log l/2a),\\ 1-(v_1/v_0)^2 &= 2\pi(\boldsymbol{e}/\boldsymbol{m})\rho a^2 v_0{}^{-2}(\tfrac{1}{2} + \log l/2a).\end{aligned}\right\} \qquad \ldots\ldots(14)$$

If our assumption is justified, both sides of (14) are small. The assumption ceases to be justified if

$$2\pi(\boldsymbol{e}/\boldsymbol{m})\rho a^2 v_0{}^{-2} = 2\pi(\boldsymbol{e}/\boldsymbol{m})j_0 a^2 v_0{}^{-3} = 2(\boldsymbol{e}/\boldsymbol{m})i\,v_0{}^{-3}$$

is not small. Here $i=\pi j_0 a^2$ is the total current.

§ 3. THE DOMAIN MODEL

3.1. In this section we think of a beam as split up into several parts. One such part we call a *domain*.

The treatment of our problem is then simplified by the following *assumptions*:—

(*a*) We consider the behaviour of a domain under the influence of its own initial distribution of charge and velocity only. All external influences, e.g. of neighbouring domains or of the boundaries, are neglected.

(*b*) We assume that initially the domain is bounded by two perpendicular cross-sections of the beam $2A$ apart. Charge density and velocity distribution at the time $t=0$ are given by

$$-\rho = -\rho_0(a),$$

$$v = v_0 + w_0(a); \quad w_0(0) = 0.$$

Here a is the distance (positive or negative) from the centre of the domain. We shall imagine that the modulation $w_0(a)$ of the mean velocity v_0 of the electrons is due to their just having passed through a modulating grid at the time $t=0$. We assume that before their passage through the grid they moved with uniform velocity v_0, so that the beam represented a steady current of density j. At the time $t=0$, the charge density of the domain will no longer be uniform, because an interval of time has elapsed since the first electrons passed the modulating grid. During this time the modulation of velocity has altered the relative position of the electrons of the domain. As a first approximation we shall, however, assume that

$$\rho_0(a) = \rho_0 = j/v_0.$$

The domain has passed the modulating grid in $2A/v_0$ seconds. Later we shall assume that the modulation is periodic with frequency f, and we shall choose A so that $2A/v_0$ is equal to a complete cycle of the modulation, i.e. $2A/v_0 = 1/f$.

(*c*) There is no mutual overtaking of electrons.

3.2. *Equation of motion.* To simplify the analysis of the model, we subtract the common translational velocity v_0 from the absolute velocity v of an electron, so that we are only studying the distribution of relative velocities $w = v - v_0$. This is equivalent to assuming that the observer is moving along the beam with velocity v_0.

As in § 2, we regard the space charge of the domain as a collection of thin condenser plates. Each one of these exerts a force proportional to its charge density on an electron. The total force on an electron at a distance a from the centre is, therefore, by the reasoning of § 2,

$$m\frac{dv}{dt} = m\frac{dw}{dt} = 2\pi e\left\{\int_{-A}^{a} \rho\, du - \int_{a}^{A} \rho\, du\right\}. \qquad \ldots\ldots(15)$$

Because of assumption (*c*), the right-hand side of (15), which represents $-2\pi e$

times the difference of charge per unit cross-section on both sides of the electron, is constant throughout the motion for a given electron.

When the initial distribution is uniform, (15) becomes

$$\frac{dw}{dt} = \frac{dw(t,a)}{dt} = 2\pi(\boldsymbol{e}/\boldsymbol{m})\rho_0 . 2a \qquad \ldots\ldots(16)$$

for an electron which is at a distance a from the centre at the time $t=0$. Integrating, we obtain

$$w(t,a) = 2\rho at + w_0(a),$$

where $\sigma = 2\pi(\boldsymbol{e}/\boldsymbol{m})\rho_0$. Let an electron originally at a be at a distance $z(t,a)$ from the centre of the domain at the time t. One further integration gives us

$$z(t,a) = a(\sigma t^2 + 1) + w_0(a)t. \qquad \ldots\ldots(17)$$

3.3. *Charge-density maxima.* By (17) the motion of the domain is completely determined. From the principle of conservation of charge, we deduce the charge density at z at the time t by means of the relation

$$\rho(z,t)\Delta z = \rho_0 \Delta a,$$

which shows that

$$\rho(z,t) = \rho_0(\partial z/\partial a)^{-1}. \qquad \ldots\ldots(18)$$

We investigate in particular the occurrence of maxima of charge density. We have to distinguish between two kinds of such maxima. The first kind is characterized by the fact that it occurs when and where overtaking of electrons occurs. For this kind we propose to retain the term *bunching*. The second kind of maximum is characterized by the fact that it occurs where electrons do not overtake each other; such a maximum is closely analogous to a region of maximum density of a longitudinally vibrating elastic rod. This type of maximum we propose to call a *compression*. The two types are illustrated in the space-time diagrams (figures 3 *a* to 3 *c*). The charge density between z and $z+\Delta z$ at the time t_1 is proportional to the number of space-time lines crossing the line $t=t_1$ between z and $z+\Delta z$.

We shall now write down the conditions for bunching and for compression. Bunching occurs when two successive layers of electrons coincide, that is to say, the co-ordinate z of two infinitesimally close electrons becomes equal. The mathematical expression for this is

$$\frac{\partial z(t_b, a)}{\partial a} = 0. \qquad \ldots\ldots(19)$$

For a given value of a, we can find the time of bunching t_b from this equation. We must bear in mind, however, that only the minimum of t_b for varying a, t_b', say, has physical significance. t_b' is the time at which bunching first occurs. For $t > t_b'$ assumption (*c*) is no longer satisfied, and, therefore, our equations no longer hold. To find t_b' we notice that

$$\left(\frac{dt_b}{da}\right)_{t=t_b'} = 0.$$

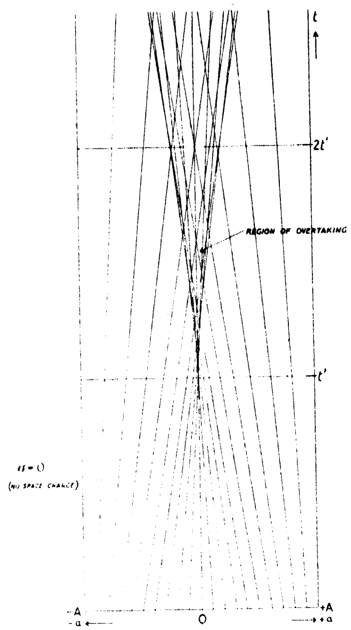


Figure 3 (*a*). Space-time diagram illustrating the no-space-charge case.

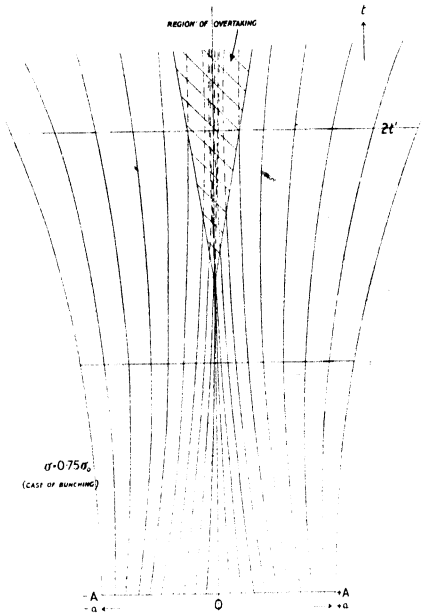


Figure 3 (*b*). Space-time diagram illustrating the case when the space-charge density is less than the critical.

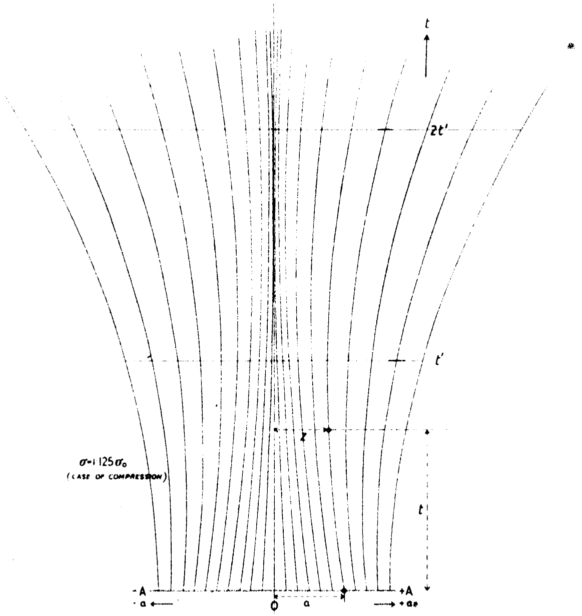


Figure 3 (*c*). Space-time diagram showing the case when the space-charge density is greater than the critical.

Differentiating (19) with respect to a we have

$$\frac{\partial^2 z}{\partial a^2} + \frac{dt_b}{da}\frac{\partial^2 z}{\partial a\,\partial t} = 0,$$

and, therefore, for $t_b = t_b'$,

$$\partial^2 z/\partial a^2 = 0. \qquad \ldots\ldots(20)$$

This is also the condition for " optimum bunching ", in the sense that electrons originally filling an interval Δa are compressed into an interval of order less than $(\Delta a)^2$ at the point of bunching. Collecting all our information, we have the following conditions for bunching:

$$\frac{\partial z}{\partial a} = 0; \quad \frac{\partial^2 z}{\partial a^2} = 0. \qquad \ldots\ldots(21)$$

We find now the condition for compression. Electrons originally between the planes $a + \Delta a$ and a are compressed after a time t into a layer of thickness Δz. To the first order of magnitude of small quantities

$$\Delta z = \frac{\partial z}{\partial a}\Delta a. \qquad \ldots\ldots(22)$$

Necessary conditions that Δz should be a minimum are

$$\partial \Delta z/\partial a = 0, \quad \partial \Delta z/\partial t = 0,$$

or, using (22),

$$\partial^2 z/\partial a^2 = 0, \quad \partial^2 z/\partial t\,\partial a = 0. \qquad \ldots\ldots(23)$$

3.4. *Choice of $w_0(a)$.* We shall now use our results to discuss two special cases obtained by putting

$$\left.\begin{aligned} &\text{(i)}\ \ w_0(a) = u(a) = v_0\{(1 - M\sin(\pi a/A))^{\frac{1}{2}} - 1\}, \\ &\text{(ii)}\ \ w_0(a) = y(a) = v_0\{(1 + M\sin(\pi a/A))^{-\frac{1}{2}} - 1\}. \end{aligned}\right\} \qquad \ldots\ldots(24)$$

$u(a)$ is the velocity distribution generated by a sinusoidal modulating voltage. $y(a)$ has the advantage of leading to expressions identical with those given in a previous paper by one of the authors (Kompfner, 1940). Both distributions are approximately equal to $-\frac{1}{2}v_0 M\sin(\pi a/A)$ for small values of M.

Neither $u(a)$ nor $y(a)$ has an exact counterpart in reality. But it is evident that discrepancy between assumed and practically realizable distributions becomes noticeable only for values of M near unity. For values of M up to 0·5, say, we can be assured of a high degree of similarity to a real velocity distribution.

3.5. *The case of no space-charge.* Before applying our formulae to the velocity distributions (24) it will be useful to discuss the motion disregarding space-charge. In this case (17) reduces to

$$z(t, a) = a + w_0(a)t. \qquad \ldots(25)$$

The condition for bunching is still given by (21). Compression cannot occur in this case. From (21) and (25) we have for the time of optimum bunching

$$t' = -(dw_0/da)^{-1} \text{ when } d^2w/da^2 = 0.$$

A simple calculation shows that in case (i)

$$t' = (\pi f M)^{-1} C_1, \qquad \ldots\ldots (26)$$

where

$$C_1 = M[2\{1 - \surd(1 - M^2)\}]^{-} .$$

In case (ii) we have

$$t' = (\pi f M)^{-1} C_2, \qquad \ldots\ldots (27)$$

where

$$C_2 = \frac{1}{\sqrt{2}}[2 - \surd(1 + 3M^2)]\,[(1 + \surd(1 + 3M^2)]^{-\frac{1}{2}}.$$

In both cases we have used the relation $2A/v_0 = 1/f$.

The constant C_2 is identical with the " correction-factor " c of the paper cited above. For small M, C_1 and C_2 are very nearly unity.

3.6. *Times of optimum bunching and compression.* We can now proceed to calculate the time of optimum bunching in the case of space-charge. By (17) we can write the condition for (optimum) bunching in the form

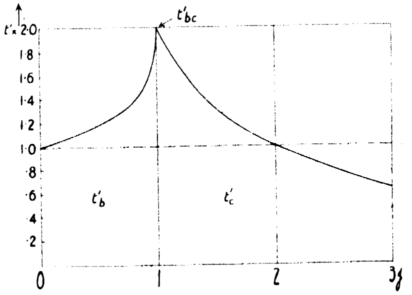


Figure 4. Time of optimum bunching and compression as function of the beam current j.

$$\frac{\partial z}{\partial a} = \sigma t^2 + 1 + \frac{dw_0(a)}{da}\,.\,t = 0 \qquad \ldots\ldots (28)$$

and

$$d^2 w_0/da^2 = 0. \qquad \ldots\ldots (29)$$

From (29) we find a. Since this equation is independent of σ and t, the solution yields the same value as in the case of no space-charge, and, therefore, substituting for a in (28) $\sigma t^2 + 1 - t/t' = 0$,

$$t_b' = \frac{1 - \surd(1 - 4\sigma t'^2)}{2\sigma t'}. \qquad \ldots\ldots (30)$$

The value of t' is given either by (26) or by (27), according to the choice of $w_0(a)$.

In the case of compression we have similarly

$$\frac{\partial^2 z}{\partial t\,\partial a} = 2\sigma t + \frac{dw_0}{da} = 2\sigma t - t'^{-1} = 0,$$

so that

$$t_c' = 1/2\sigma t'. \qquad \ldots\ldots (31)$$

If $4\sigma t'^2 < 1$, t'_b is real, and bunching occurs:—If $4\sigma t'^2 > 1$ the expression (30) is complex, bunching cannot occur and we shall have compression instead. For

$$\sigma = \sigma_0 = \tfrac{1}{4} t'^{-2}$$

we have a border-line case between bunching and compression; some electrons catch up with those immediately in front of them, but they have not sufficient speed to overtake them. The time at which this catching up occurs is given by using this value of σ in (31); thus,

$$t'_{bc} = 2t'.$$

If $\sigma = \sigma_0$, we find from the equations

$$\sigma = 2\pi(\boldsymbol{e}/\boldsymbol{m})\rho, \quad j = \rho v_0,$$

that the current-density j is given by

$$j = \frac{v_0}{2\pi(\boldsymbol{e}/\boldsymbol{m}) \,.\, 4t'^2},$$

or, using (26) and (27),

$$j = \frac{\pi v_0 M^2 f^2}{8(\boldsymbol{e}/\boldsymbol{m}) C_{1,2}^2} = 1{\cdot}46 \,.\, 10^{-20} \sqrt{V_0} f^2 M^2 C_{1,2}^{-2} \text{ amp./cm.}^{-2}.$$

where V_0 is the initial accelerating potential difference in volts.*

For current densities below the critical density j we have bunching proper, while at current densities above j we have compression only.

3.7. *Compression factor.* By the term *definition* we understand the minimum thickness Δz into which a given portion of initial thickness Δa, called the *aperture*, can be compressed. The ratio aperture/definition $= \Delta a/\Delta z$ is called the *compression factor*. In the case of the chopped beam particularly, but also quite generally, it will be of interest to investigate how the compression factor depends on the aperture and the choice of various other factors.

We write $\Delta w_0(a)$ for $w_0(a + \Delta a)$. By a simple calculation we find from (17) that the minimum Δz is

$$\Delta a \left[1 - \frac{1}{4\sigma} \left(\frac{\Delta w_0}{\Delta a} \right)^2 \right],$$

and that it is attained at the time

$$t_m = -\frac{1}{2\sigma} \frac{\Delta w_0}{\Delta a}.$$

In the case of compression, this formula is true without restriction. In the case of bunching it is only valid if $t < t_b'$, since after that time our formulae are no longer applicable.

If we restrict ourselves to small M with

$$w_0(a) = -\tfrac{1}{2} v_0 M \sin(\pi a/A),$$

and to apertures symmetrical about $a = 0$, the compression factor becomes

$$\frac{\Delta a}{\Delta z} = \left[1 - \frac{1}{4\sigma} \left(\frac{\pi v_0 M \sin \Delta x}{2A\Delta x} \right)^2 \right]^{-1},$$

---

* Thus in a practical example where $V_0 = 3600$ volts, $\lambda = 8$ cm. ($f = 3{\cdot}75 \times 10^9$ cycles/sec.) and $M = 1/\pi$, the critical current j, above which no proper bunching would be possible, will be $1{\cdot}4$ amp./cm.$^{-2}$. For $\lambda = 1$ cm., j will be of the order of 100 amp./cm.$^{-2}$.

where $\Delta x = \pi\Delta a/2A$. If $2A = v_0/f$, $\pi fM = t'^{-1}$, $\sigma_0 = \frac{1}{4}t'^{-2}$,

$$\frac{\Delta a}{\Delta z} = \left[1 - \frac{\sigma_0}{\sigma}\left(\frac{\sin \Delta x}{\Delta x}\right)^2\right]^{-1}.$$

It can be seen from this equation and from the graph (figure 5) that the quality of the compression deteriorates rapidly as the current density increases, while the time of optimum compression diminishes.

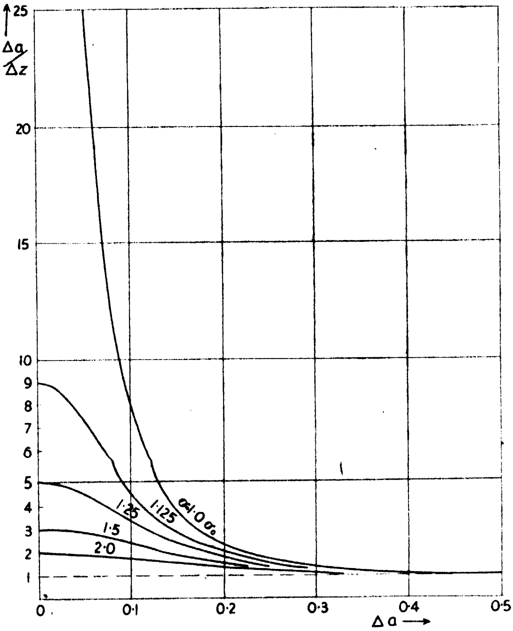


Figure 5. Compression factor $\Delta a/\Delta z$ as function of aperture Δa ($M \ll 1$).

3.8. *Validity of results.* The domain model has the obvious disadvantage that its assumptions are highly artificial. Even so, we believe that at least the following qualitative result of this section is correct:—

There will be overtaking and proper bunching with attendant split-bunching when the current density is below a certain critical value. Above this critical value there will be no proper bunching, the mutual repulsion of electrons will prevent any overtaking. If the current density is taken too high, the current peaks will become ***low and flat.***

We can even argue that the numerical results of this section have some practical significance for the following two reasons:—

1. The error introduced by assuming the cross-section of the beam as infinite will be relatively small if the domain is not too wide. It may also be partially offset by the total neglect of the space-charge outside the domain.

2. In the middle of the domain, the effects of the space-charge on both sides outside the domain will probably cancel each other. To a first approximation, the effect of the neglected space-charge will be a compression of the domain, preventing the gradual broadening which takes place in our model.

If these two assumptions are granted, it follows, for example, that the formulae for the time of optimum bunching and compression must have a fairly close relation to reality.

Advantages of the domain model are its great simplicity, and the fact that its application does not in any way presuppose that the modulation is small and sinusoidal.

§4. ACKNOWLEDGEMENT

We wish to thank Dr. G. Liebmann for much helpful criticism.

REFERENCES

FAY, SAMUEL and SHOCKLEY, 1938. *Bell Syst. Tech. J.* **17**, 49.
KOMPFNER, 1940. " Velocity Modulation ", *Wireless Engr*, Nov.
PLATO, KEEN and ROTHE, 1936. *Z. Phys.* **101**, 509.
RAMO, 1939. *Proc. Inst. Rad. Engrs*, 756.
WEBSTER, D. L., 1939. " Cathode Ray Bunching ", *J. Appl. Phys.* **10**, 501.

Note added 20. 12. 41. At the time of writing of this paper we had no access to scientific literature; we have found since that the problem of a steady flow of electrons between infinite plane electrodes has been treated several times before, and that the existence of the " potential hill " and the resulting formula for maximum current are well known. Very exhaustive investigations of unmodulated electron beams are to be found in the papers by Fay, Samuel and Shockley (1938), and by Plato, Keen and Rothe (1936).

The problem of velocity-modulated beams has also been treated by several writers besides Webster. A most interesting theory of beams of circular cross-section focussed by means of an " infinite " axial magnetic field has been developed by Hahn and Ramo (see Ramo, 1939). Under the assumption of a small modulation they prove the existence of " electronic waves ". In our terminology these waves could be described roughly as compressions travelling along the beam at a velocity slightly different from the mean velocity of the beam. The " electronic wave theory " is much superior to the elementary theory developed in this paper, in so far as it takes account of the finite cross-section and of the influence of the screening boundary of the drift space. It should be pointed out, though, that the boundary conditions assumed by Ramo (infinite length of beam, variation of electron velocity with distance from the axis) are not satisfied in reality, so that the results can only be regarded as qualitative. Since our discussion of velocity-modulated beams is physically more intuitive and mathematically much simpler than Ramo's, we hope that the publication of the present paper may still serve a useful purpose.

FUNDAMENTAL LAWS AND DEFINITIONS IN PHYSICS: I—OHM'S LAW *

By C. W. HANSEL,
Bedford School

Received 14 *November* 1941

§ 1. INTRODUCTION

THERE are some sections of Physics the treatment of which may be varied according to taste or opinion. Much harm might be done in restricting the scope of the discussion to any particular method of treatment or to any single viewpoint. There are other matters in which it seems highly desirable to adopt greater uniformity of treatment and point of view, and a standardization of notation and terms is highly desirable.

The object of this article is to stimulate co-operative discussion with a view to discovering the best elementary treatment of Ohm's law.

§ 2. OHM'S LAW

What did Ohm discover? Is the most useful introduction to Ohm's law a study of Ohm's actual investigations? What is the most useful statement of the law?

An account of the experimental work of Ohm is given in *A History of Physics* by Florian Cajori. He first established the laws relating to the resistance of metallic wires embodied in the formula, $r = \frac{l}{a}S$. He experienced difficulties in these experiments owing to variations of his batteries, and he finally used thermoelectric elements as his sources of E.M.F. He then measured the strength of the current in a circuit, keeping the E.M.F. constant and varying the resistance of a portion of it. He found results in agreement with the formula

$$X = \frac{a}{b+x},$$

where X is proportional to the current in a conductor forming part of the circuit whose length x was varied. a and b are constants.

---

* Paper to be read at a joint meeting with the Science Masters' Association, 9 April 1942.

By repeating his experiments with a different E.M.F. he showed that *a* assumed a new value but not *b*. Thus he identified *a* with the E.M.F., and *b* with the resistance of that part of the circuit which he did not vary.

Thus the formula $X = \frac{a}{b+x}$ is equivalent to

$$\text{Current} = \frac{\text{E.M.F.}}{\text{Resistance}} . \quad \ldots\ldots\ldots \quad 1$$

This work was published in 1826.

In 1827, Ohm gave a theoretical investigation of electrical conductivity in metallic conductors based on Fourier's mathematical treatment of the conduction of heat. This is referred to in the *Encyclopaedia Britannica*, Article ***Ohm***, as though it constituted the whole of his work. " It may be doubted whether Ohm's investigation could have been made but for the magnificent work of Fourier on the Conduction of Heat." It will be observed that the electric conductivity of a metal was proved by Ohm to be independent of the potential difference, but the thermal conductivity was incorrectly assumed by Fourier to be independent of temperature.

The Ninth Edition of the *Encyclopaedia Britannica* contains a more appreciative account of the work of Ohm by Professor George Chrystal, but no mention is made of his experimental work. In the article on ***Electricity*** he says: " The theory on which Ohm based his law may be and has been disputed, but the law itself and the applications which Ohm and others have made of it are in the fullest agreement with all the known facts." Later in the same article, the equivalent of the following formula is given:—

$$\frac{E}{I} = r = \frac{ls}{a},$$

where l, a, s, relate to a metallic wire " and s is a constant depending on its material, temperature, and physical condition generally. This is Ohm's law ". 2

Ohm's actual investigations do not appear to provide a suitable introductory treatment to Ohm's law.

A committee of the British Association, consisting of Professors Clerk Maxwell, J. D. Everett, and Dr. A. Schuster, prepared a report on the testing of Ohm's law. The experimental work was carried out in the Cavendish Laboratory by G. Chrystal under the supervision of Clerk Maxwell. Quoting from the B.A. Report of 1876:

" The statement of Ohm's law is that, for a conductor in a given state, the electromotive force is proportional to the current produced ". 3

" The quotient of the numerical value of the electromotive force divided by the numerical value of the current is defined as the resistance of the conductor;

and Ohm's law asserts that the resistance, as thus defined, does not vary with the strength of the current." 4

Later, in the same report, Maxwell states:

" Ohm's law may be stated thus:—The electromotive force which must act on a homogeneous conductor in order to maintain a given steady current through it, is numerically equal to the product of the resistance of the conductor into the strength of the current through it. 5

" If, therefore, we define the resistance of a conductor as the ratio of the numerical value of the electromotive force to the numerical value of the strength of the current, Ohm's law asserts that this ratio is constant—that is, that its value does not depend on that of the electromotive force or of the current.

" The resistance, as thus defined, depends on the nature and form of the conductor, and on its physical condition as regards temperature, strain, etc., but if Ohm's law is true it does not depend on the strength of the current."

The Dictionary of Physics states the results of these experiments as showing that " the resistance of a conductor one square centimetre in section is not diminished by as much as $1/10^{12}$ of its value—assuming the temperature constant —when a current of 1 ampere traverses it. We may consider, therefore, that Ohm's law is verified and that the ratio of the E.M.F. to the current is constant ".

" Ohm's law states that in any circuit under constant physical conditions the ratio of the electromotive force to the current is a constant. This constant is known as the resistance of the circuit; its reciprocal is the conductivity."

. 6

§ 3. STATEMENTS OF OHM'S LAW

J. J. Thomson, *Elements of Electricity and Magnetism*, 1904.

" Ohm's Law. The relation between the electromotive force and the current was enunciated by Ohm in 1827, and goes by the name of Ohm's Law.

" This law states that if E is the electromotive force between two points A and B of a wire, I the current passing along the wire between these two points, then

$$E = RI, \qquad . . . . 7$$

where R is a quantity called the resistance of the wire ".

" The point of Ohm's law is that the quantity R defined by this equation is independent of the strength of the current flowing through the wire, and depends only upon the shape and size of the wire, the material of which it is made, and upon its temperature and state of strain.

" The most searching investigations have been made as to the truth of this law when currents pass through metals or electrolytes; these have all failed to discover any exceptions to it, though from the accuracy with which resistance can be measured (in several investigations an accuracy of one part in 100,000 has been attained) the tests to whch it has been subjected are exceptionally severe."

W. C. D. Whetham, *The Theory of Experimental Electricity*, 1923.

" Ohm's law, that the current c varies as the applied electromotive force E, may be stated in the form

$$c = kE. \quad \dots \dots \dots \quad 8$$

where k is a constant known as the conductivity of the conductor.

" Another mode of statement, more usually adopted, is to say that

$$c = \frac{1}{R} E \quad \text{or} \quad E = Rc,$$

where R is a constant which, like the conductivity, depends only on the nature, dimensions and temperature of the conductor, and is independent of the current. This constant R is called the resistance."

R. W. Pohl, *Physical Principles of Electricity and Magnetism*, translated by Winifred M. Deans, 1930.

" In general we define the resistance r of any conductor to be the quotient

$$\frac{\text{Pressure } (E) \text{ between the ends of the conductor}}{\text{Current } (i) \text{ in the conductor}}.$$

" This quotient E/i, which is defined to be the resistance r, is not in general constant for the conductor in question. The resistance r depends in a complicated way on the current i and on numerous conditions of experiment. It is only in particular cases that the quotient E/i is constant. In these particular cases Ohm's law—which is one of our most important experimental results—is true. In words:

" The pressure E between the ends of a conductor is proportional to the current i in the conductor, or

$$E = iR. \quad \dots \dots \dots \quad 9$$

" A metallic conductor at constant temperature is an example of a particular case where Ohm's law holds."

Alexander Russell, *Alternating Currents*, 1904 and 1914.

" Ohm proved experimentally that the ratio of the difference of potential to the current was constant so long as the conductor between the two points remained in the same physical state. This constant ratio is called the resistance of the conductor between the points, and Ohm's law may be written

$$i = \frac{e}{r}\text{"} \quad \dots \dots \dots \quad 10$$

J. H. Jeans, *Electricity and Magnetism*, 1923.

" Ohm's Law. The difference of potential between any two points of a wire or other linear conductor in which a current is flowing, stands to the current flowing through the conductor in a constant ratio, which is called the resistance between the two points. 11

"It is here assumed that there is no junction with other conductors between these two points, so that the current through the conductor is a definite quantity."

§4. E.M.F. AND POTENTIAL DIFFERENCE

In Whetham's *Electricity* the following passage occurs:—"In the science of current-electricity, it is usual to call a difference of potential an electromotive force. The name is not a happy one, for an electric force f has the physical dimensions [force/quantity of electricity], and is related to the potential-difference V, or [work/quantity of electricity], by the equation $f = -dV/dx$. The use of the name electromotive force as a synonym for difference of potential, however, is established firmly, ..."

The statements of Ohm's law numbered 1 to 8 all use the term E.M.F. synonymously with potential difference. Statements 9 to 11 employ the term potential difference and do not refer to E.M.F. The term potential difference or "pressure drop" is favoured by electrical engineers.

Some writers distinguish between the two terms, and whilst there is no advantage in retaining the use of two terms with the same meaning, there seems to be an advantage in restricting the use of E.M.F. to the voltage produced by a generator. Only a battery, dynamo, induction coil, and the like, can have an E.M.F., but anything through which a current of electricity is flowing has a potential difference equal to the strength of the current flowing multiplied by the resistance of the conductor. This potential difference is provided by the generator in the circuit. This is merely a statement of Kirchhoff's second law relating to a current-carrying network of conductors. Again, in any non-generator, the current flows from a high to a low potential, but in a generator there is a transformation of energy resulting in the electricity rising in potential in the direction of current flow. The rise in potential equivalent to the transformation of energy per unit quantity of electricity is the E.M.F. (due allowance being made for the natural fall in the potential due to resistance of the generator).

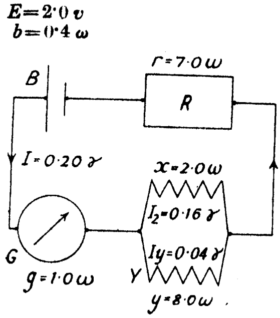
The use of the term E.M.F. with two meanings is undesirable. The E.M.F. of a cell always means the potential difference at its terminals when on open circuit and not the potential difference for the cell when it is sending a current.

§5. NOTATION

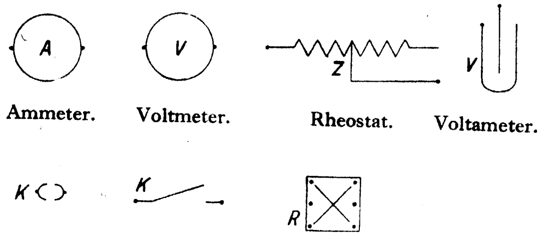
A standard notation is desirable in symbols, diagrams, and working. The writer uses the notation indicated in the diagram (figure 1).

A suitable Greek letter may be preferable to v for volts.

i, *e*, *v*, *r*, could be used for current, E.M.F., potential difference, and resistance in electromagnetic units (e.m.u.), but *v* and *r* are seldom required.



$$R=b+r+g+x,y=10{\cdot}0\ \omega,$$
$$I=\frac{E}{R}=0{\cdot}20\ a,$$
$$V_b=Ib=0{\cdot}08\ v,$$
$$V_r=Ir=1{\cdot}40\ v,$$
$$V_g=Ig=0{\cdot}20\ v,$$
$$V_{x,y}=I(x,y)=\underline{0{\cdot}32\ v},$$
$$E=\underline{2{\cdot}00\ v},$$
$$I_x=\frac{V_x}{x}=\frac{V_{x,y}}{x}=0{\cdot}16\ a,$$
$$I_y=\frac{V_y}{y}=\frac{V_{x,y}}{y}=0{\cdot}04\ a.$$
$$I=\underline{\overline{0{\cdot}20\ a.}}$$



Ammeter. Voltmeter. Rheostat. Voltameter.

Make and break key. Current reverser.

Figure 1.

Resistances are written external to the mesh and currents within it.

b, *r*, *g*, *x*, *y*, are the numerical values of the resistances of the circuit units B, R, G, X, Y, respectively.

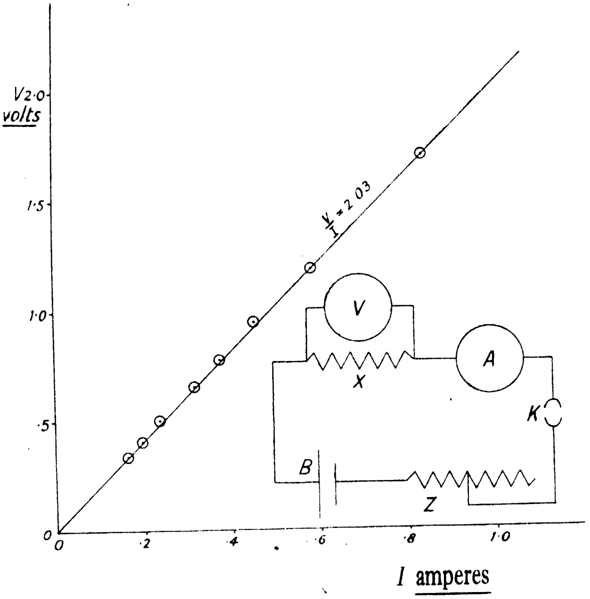
§ 6. ILLUSTRATION OF OHM'S LAW

Figure 2 shows a form of apparatus suitable for beginners. The object of the experiment is to apply a series of different potential differences (V) to a metallic conductor X (a 2-ohm coil is suitable), and to observe how the current (I) through the conductor varies with the applied potential difference.

The ammeter A (range 0 to 1·0 ampere), may be calibrated with a tangent galvanometer, but this does not form part of the experiment, and a discussion of this will only distract the attention of the beginner from the main issue. Similarly, the high resistance voltmeter (V) (range 0 to 2·5 volts, resistance 500 ohms or more) may be calibrated in any way whatever, but in this experiment the scale is taken as correct.

The key K is an important part of the apparatus. The readings should be taken as quickly as is consistent with reasonable accuracy, and the current should only be allowed to flow for short periods of time to avoid the coil X being heated by the current. It is important to direct attention to this matter, although with a manganin coil the temperature effect is well within the error of reading even if little attention is paid to this matter.

Figure 2 shows results obtained without special care and by a beginner.



V	I	$\frac{V}{I}$	$\frac{I}{V}$
1·72	0·85	2·02	0·495
1·20	0·59	2·03	0·493
0·95	0·46	2·06	0·485
0·78	0·38	2·05	0·488
0·65	0·32	2·03	0·493
0·50	0·24	2·08	0·481
0·40	0·20	2·00	0·500
0·34	0·17	2·00	0·500

Figure 2.

The low readings are accidentally more accurate than the higher readings, or it may be a compensation of scale and experimental error.

If the potential difference be regarded as the independent variable, I is plotted against V, and the graph obtained shows that I/V is constant or $I = kV$, where k is the conductance of X.

If the current be regarded as the independent variable, V is plotted against I, and the graph (figure 2) shows that V/I is constant or $V = xI$, where x is the resistance of X.

Thus the current through a metallic conductor under fixed physical conditions is directly proportional to the potential difference applied to it. This is Ohm's law.

If Ohm's law holds for every part of a circuit, then the current in a circuit whose E.M.F. is varied is directly proportional to the E.M.F. In this case $E = RI$, and this same relation holds in such a circuit if E is constant and R is varied. In a circuit whose E.M.F. is due to thermo-elements, or if the circuit contains a variable-speed dynamo with constant separately-excited field, the E.M.F. may be varied. In these circuits R is not definitely constant because the temperature is not constant throughout the whole of the thermoelectric circuit, and the resistance of a brush contact of a dynamo varies with the current density over the contact area and with the peripheral speed of the commutator. Similarly, in a battery circuit in which the E.M.F. is constant, the relation $E=RI$ is not exactly true because the resistance of a cell varies somewhat with the current that it sends through the circuit. Thus it is not correct to say that for a complete circuit

$$\text{Current} = \frac{\text{E.M.F.}}{\text{Total resistance}}.$$

§ 7. UNITS

It may be advisable to distinguish between the definition and the description of a unit. The definition of a unit should be framed so that the relation of the unit to the fundamental units of length, mass, and time is clearly indicated.

Definition. An ampere is 1/10 of that current which flowing along a conductor 1 centimetre in length bent into the arc of a circle of 1 centimetre radius produces a force of 1 dyne on a unit pole at the centre of the circle.

The dimensions of current may be found from this definition. In addition the formula $f=2\pi In/10r$ immediately follows from it if the inverse-square law for the variation of magnetic force with distance be assumed. This law was verified experimentally by Maxwell.

What is called the legal or practical definition of the ampere is no definition at all, any more than the definition of a shilling is that it is equal to 12 pence. It is one description of an arbitrarily chosen chemical effect of the current. The mass of silver deposited by one ampere in one second or by one coulomb could not be ascertained until the ampere had been defined, and the value found by experiment can never be accepted as entirely accurate. In what proportions are the isotopes of silver deposited during electrolysis? Is the equivalent weight of silver determined by chemical analysis exactly the same as that found by electro-deposition?

Why is it necessary to have two units? Why not make the ampere 10 times larger?

Definition. A volt is the potential difference between two points if a joule of work is done in transferring a coulomb from one point to the other.

The dimensions of potential difference may be found from this definition. Also, assuming the law of Conservation of Energy, the formula $JH=VIt$ is immediately obtained. Again, for a conductor that obeys Ohm's law it follows that $JH=I^2Rt$ (verified experimentally by Joule).

The practical " definition " of the volt in terms of the E.M.F. of a normal Weston cadmium cell is one description of many that could be given of the volt. A volt is also the E.M.F. induced in a conductor when it cuts magnetic lines of force at the rate of 10^8 per second.

Definition. A conductor has a resistance of 1 ohm if a current of 1 ampere flows through it when a potential difference of 1 volt is applied to it.

The practical " definition " of the ohm is far from being practical. A tube of uniform cross section of 1 square millimetre 106·3 centimetres in length is impossible of realization, and what is the isotopic composition of pure mercury?

§8 CONCLUSIONS

1. It is desirable to standardize terms, definitions, symbols, and diagrams.

2. It is desirable to assign different meanings to the terms E.M.F. and potential difference.

3. Ohm's law applies most accurately to metallic conductors.

For a metallic conductor under fixed physical conditions, the ratio of the potential difference applied to it to the current flowing through it is constant. This is Ohm's law.

4. The resistance of a conductor is the ratio of potential difference applied to it to the current passing through it.

The resistance of a conductor is not, in general, constant.

Conductance is the reciprocal of resistance.

5. Is it necessary to have

$$\begin{aligned} 1 \text{ ampere} &= 10^{-1} \text{ e.m.u.} \\ 1 \text{ volt} &= 10^{8} \text{ e.m.u.} \\ 1 \text{ ohm} &= 10^{9} \text{ e.m.u. ?} \end{aligned}$$

Could the ampere be made equal to 1 e.m.u., and if the volt and ohm are not equated to 1 e.m.u. could each be made 10^7 e.m.u.?

6. The practical definitions are not suitable for the theoretical development of the subject.

FUNDAMENTAL LAWS AND DEFINITIONS IN PHYSICS: II—SPECIFIC HEAT AND NEWTON'S LAW OF COOLING*

By C. W. HANSEL,
Bedford School

Received 9 *December* 1941

§1. INTRODUCTION

THE terms thermal capacity and specific heat are not always defined in the same way. For example:—

J. Clerk Maxwell, *Theory of Heat*, 1871 and 1894.

" The capacity of a body for heat is the number of units of heat required to raise that body one degree of temperature." 1

---

* Paper to be read at a joint meeting with the Science Masters' Association 9 April 1942.

" The Specific Heat of a body is the ratio of the quantity of heat required to raise that body one degree to the quantity required to raise an equal weight of water one degree." 2

Sir William Thomson, *Encyclopaedia Britannica* (9th edition). Article *Heat.*

" Definition 1. The thermal capacity of a body . . . . is the quantity of heat required to raise its temperature by one degree on the absolute thermodynamic scale." 3

" Definition 2. The specific heat of a substance is the thermal capacity of a stated quantity of it.

This stated quantity is generally understood to be the unit of mass, ..." 4

H. L. Callendar, *Encyclopaedia Britannica* (10th and subsequent editions). Article *Calorimetry.*

" The thermal capacity of a body is measured by the quantity of heat required to raise its temperature one degree, ..." 5

" The specific heat of a substance is sometimes defined as the thermal capacity of unit mass, but more often as the ratio of the thermal capacity of unit mass of the substance to that of unit mass of water at some standard temperature. The two definitions are identical, provided that the thermal capacity of unit mass of water, at a standard temperature, is taken as the unit of heat. But the specific heat of water is often stated in terms of other units." 6

Thomas Preston, *Theory of Heat*, 1904.

" The thermal capacity of a body is defined as the quantity of heat necessary to raise the temperature of the body 1° c., and the thermal capacity of a substance is the quantity of heat required to raise unit weight (one gramme) of the substance 1° c." 7

" The specific heat of a substance is its thermal capacity compared with that of water; . . . ." 8

J. H. Poynting and J. J. Thomson, *Heat*, 1906.

" The specific heat of a substance is the number of calories needed to raise 1 gramme of the substance 1° c." 9

Capacity for heat of a calorimeter is mentioned but not emphasized.

A. W. Barton, *Heat*, 1933.

" The thermal capacity of a body is the amount of heat to raise its temperature through 1° c. 10

" The specific heat of a substance is the amount of heat needed to raise the temperature of 1 gramme of it through 1° c." 11

§ 2. THERMAL CAPACITY AND SPECIFIC HEAT

In the statements 1 to 11 there are various definitions of thermal capacity, but are they of any practical importance ?

It is scarcely necessary for a beginner to know this term, and it might be omitted from examination questions.

The water equivalent of a calorimeter is not quite numerically the same as the thermal capacity of it since it is a mass of water which in an actual experiment and under actual working conditions would absorb the same amount of heat as the calorimeter absorbs. This is not quite the same as the mass of the calorimeter multiplied by its specific heat.

The term specific heat is sometimes defined as a ratio and sometimes as a quantity of heat per unit mass per unit rise of temperature. If specific heat is to enter into an equation of the form $H = ms(t_2^\circ - t_1^\circ)$, it cannot be a pure number, and for this reason I think the ratio definition should be abandoned.

Again, specific heat is the heat per unit mass (gram, pound, or kilogram) per degree (any). In this case consistent units must be used in the heat equation.

The British Thermal Unit of Heat involves the pound as the unit of mass. The heat engineer borrows this unit and uses it instead of the slug-degree-Fahrenheit unit. In Joule's determination of the mechanical equivalent of heat, the heat is in British Thermal Units and the work is in foot-pounds.

I favour the definition of specific heat given in Chambers' *Dictionary* :—

" Specific Heat, the number of heat-units necessary to raise the unit of mass of a given substance one degree in temperature."

This definition is dimensionally correct and applies to all consistent sets of units.

§ 3. NEWTON'S LAW OF COOLING

F. H. Schofield, *Dictionary of Applied Physics* (*Heat*, *Convection of*).

" Newton propounded a law to the effect that the rate of cooling of a hot body in a stream of air is proportional to the difference in temperature between the body and the air." 1

And in a footnote, " This law is often taken to apply to natural convection, though Newton expressly said that it was given for a body ' not in still air but in a uniform current of air.' "

" The heat loss by forced convection from a hot surface is proportional to the temperature difference between the surface and the ambient fluid. This has been shown by Boussinesq from hydrodynamical reasoning, by Rayleigh from the principle of similitude, and it is confirmed by a considerable mass of experimental evidence."

" Conclusions on natural convection. (1) The heat loss from a hot surface is approximately proportional to $\theta^{5/4}$, where θ° is the temperature difference between the hot surface and the ambient fluid."

See also the article on *Dynamical Similarity* in the same volume by Hyman Levy.

Engineers appear to have applied Newton's law in the form stated by Newton, that is, to a body cooling in a draught. Newton's law is a law of forced convection, and appears to hold up to high temperature if the hot body moves in a rapidly moving current of gas. The cylinder of a moving motor-cycle cools according to this law.

Textbooks of physics often state Newton's law of cooling as follows:—

" The rate of cooling of a body is proportional to its excess temperature above the temperature of its surroundings provided that this excess temperature is small." 2

According to this statement, Newton's law is not a " law " at all. All that is asserted in this statement is that a short enough part of the cooling curve at the origin is straight. This is necessarily true whatever the shape of the cooling curve.

Still worse is that form of demonstration which reconciles Stefan's fourth-power law of radiation with Newton's law of forced convection. Surely, a black body cooling in a vacuum does not cool in the same way as an ordinary body cooling in a stream of gas.

It has been stated that an ordinary thermometer heated to 20° or 30° above the temperature of the surrounding air cools according to statement 2. The following results obtained with cooling thermometers do not confirm this. The data given refer to thermometers cooling in air at temperature 20° c.

The bulb A was blackened over burning camphor. B was the same thermometer as A, but with a bright bulb (black rubbed off).

Time (min.)	0·0	0·5	1·0	1·5	2·0	2·5
A ° c.	105·5	76·4	59·3	48·0	40·0	34·7
B ° c.	110·0	80·0	61·2	49·2	40·8	35·5
Time (min.)	3·0	3·5	4·0	4·5	5·0	
A ° c.	31·2	28·8	27·2	26·0	25·1	
B ° c.	31·8	29·3	27·4	26·3	25·4	

§ 4. THE COOLING CALORIMETER

For natural convection, the heat loss from a surface is proportional to the excess temperature to the power of 5/4. Does natural convection occur when a calorimeter is used as in a determination of specific heat by the method of mixtures? What difference of temperature is necessary between a body and surroundings for natural convection to occur? Does cooling occur according to statement 2? Is there a surface film of gas on the calorimeter which is almost at the same temperature as the calorimeter and impedes both convection and radiation and transference of heat from the calorimeter to the surrounding air?

Numerical data relating to these problems would be very valuable.

Can Newton's law of cooling be illustrated with a cooling calorimeter? Statement 2, which is not Newton's law, may be approximately applicable. Since there is no forced convection, Newton's law cannot be illustrated in this way.

Examination questions

A candidate should not be asked to verify Newton's law of cooling for a cooling calorimeter.

Even in the School Certificate Examination, experiments have been set in which cooling observations provide the data from which a graph is required of rate of cooling plotted against excess temperature. It is impossible to obtain the rates of cooling with sufficient accuracy by drawing tangents to a cooling curve. In experiments of this kind, is the plot of rate of cooling against excess temperature expected to be a straight line?

The experimental data given below throw some light on the difficulties of work of this kind and the impossibility of obtaining sufficient accuracy in the measurement of temperatures taken carefully with a thermometer graduated in tenths of a degree.

100 grams of water contained in a calorimeter of water equivalent 5 grams cooled in surroundings kept at a constant temperature $t = 16{\cdot}8^\circ$ c. The temperature of the water was t° c. at time T minutes.

T min.	0	1	2	3	4	5	6	7
t° c.	47·7	46·8	46·0	45·2	44·5	43·8	43·2	42·5
t° c. (corrected)	47·56	46·77	46·01	45·27	44·54	43·84	43·15	42·48

T min.	8	9	10	11	12	13	14
t° c.	41·8	41·1	40·5	40·0	39·4	39·0	38·4
t° c. (corrected)	41·82	41·19	40·57	39·96	39·37	38·80	38·24

From the plot of $\log (t - 16{\cdot}8)$ against T it is found that

$$\log (t - 16{\cdot}8) = -0{\cdot}0112T + 1{\cdot}488.$$

The corrected values of t have been calculated from this equation.

Although the uncorrected values of t are as accurate as could be observed to the nearest tenth of a degree, they are not accurate enough to enable the rate of cooling to be found with sufficient accuracy to show that a linear relation exists between the rate of cooling and $(t - 16{\cdot}8)$. Values found by drawing tangents to the cooling curve give even worse results. It may be thought that the temperatures might have been estimated to $0^\circ{\cdot}01$ c. This cannot be done owing to the sticking of the mercury in the thermometer.

According to these results the rate of cooling is proportional to the excess temperature to the power of 1·03.

§ 5. RECOMMENDATIONS

1. That the term thermal capacity be omitted from examination syllabuses and that examiners avoid the term when setting questions.

2. That specific heat be defined as a quantity of heat per unit mass per unit rise of temperature.

3. That Newton's law of cooling shall only be deemed to refer to a body cooling in a stream of gas. The name of the law should not be used to mean the proportionality of rate of cooling and excess temperature for a body cooling in any arbitrary way.

FUNDAMENTAL LAWS AND DEFINITIONS IN PHYSICS: III—MASS *

By C. W. HANSEL,
Bedford School

Received 25 *November* 1941; *in revised form* 31 *December* 1941

§ 1. INTRODUCTION

THREE fundamental conceptions are those of space, time and matter. They form the background of daily experience. Space, time and matter are measured in terms of arbitrary units of length, time and mass.

It is a mistake to postpone the teaching of ideas on mass on the score of difficulty (see Quotation 94 and the B.A. *Report*). The difficulty is not one of *ideas* but of *words*. A child becomes familiar with the popular use of the word "weight" in the sense of "mass" and uses the word *weight* for *mass*. What does a child know about gravitational attraction? To what extent is a child interested in the gravitational pull on a bag of sweets? The child is very definitely interested in the mass of the material, the quantity of delectable substance, not in the gravitational pull on it. The beginner in science must be trained to use his words carefully and correctly. The same word should not be capable of two different scientific interpretations. No slovenliness of record or speech should be tolerated. A similar difficulty occurs in the use of the words *heat* and *temperature*. A beginner uses the word *heat* in the popular sense of *temperature*. Persistent correction ultimately eradicates the habit.

The distinction between mass and weight should be taught early in the science course. The contention that the idea of weight is familiar and the idea of mass obscure is an incorrect way of stating that popularly the word *weight* is used in many different senses (see later) and usually in the same sense as the scientific term *mass*. Actually, this initial misuse of a word may be turned to instructional advantage if the distinction between mass and weight is emphasized and the correct use of these terms is insisted upon. A beginner has no difficulty in grasping Newton's idea of gravitational force between objects. The weight of a body is defined to be the gravitational force of the earth on it, and varies from place to place. A beginner may be led to predict that a body will have no weight at the centre of the earth. It is easily realized that a spring balance directly measures weight.

---

* **Paper to be read at a joint meeting with the Science Masters' Association, 9 April 1942.**

It is important in a science course to use a beam balance early. It is not necessary to know the theory of a beam balance before using it to measure mass, any more than it is necessary to know the construction and theory of a clock before using it to measure time. The young beginner is quite capable of understanding that the result of weighing on a beam balance is the same everywhere, and hence it does not measure weight, which is not the same everywhere. A young child has usually played with a toy set of scales and " weights ", and in any case he knows that a pair of scales is not used to measure gravitational attraction, but to measure the quantity of all kinds of different materials. Quantity of matter is mass—the beam balance measures mass.

From now on, all the results of weighing with a beam balance are recorded and spoken of as mass, and results of weighing with a spring balance as weight. If in all lessons the correct terms are used, the habitual careful and correct use of the terms mass and weight becomes established. The science teacher has laid the foundation for future extension of the work.

A few teachers of mechanics ignore the term *mass* completely. It will be generally agreed that this is undesirable. What is even more undesirable is for the teacher of mechanics to mention the *term* but not to use the *idea*. It is unfortunate that teachers of science, mathematics and engineering do not always collaborate to the fullest extent to give to each other the greatest possible assistance in developing and using this important and fundamental idea of mass. It is most desirable that a uniform treatment and a uniform system of symbols and formulae should be used by all who are concerned in the teaching of mass and related subject-matter. It is essential that all those engaged in scientific instruction should abide by the correct use and meaning of scientific terms and train the pupil accordingly. By the time the student begins the study of dynamics he should have established the habit of using the terms mass and weight carefully and correctly. The habitual correct use of the term mass may now be enlarged and used in connection with force, momentum, kinetic energy and, at a later stage, moment of inertia.

§ 2. SYSTEMS OF UNITS

Some teachers of mathematics and engineering use only one system of units. Some use three systems and a fourth hybrid " system ", more accurately described as chaos than system (see *A First Dynamics*, by C. S. Jackson and W. M. Roberts, 1909, p. 311). It will be generally agreed that a student should not only be able to use any system of units, but he should use them freely and with equal facility. In the interests of sound training, and to avoid confusion and vagueness, any so-called *fourth system using a mixture of units from more* *than one system, or using two units for the same quantity, or one name for two units* (for example, gram—gram weight—dyne), *should be rigorously excluded.*

No one should claim that all discussion should be in terms of gravitational units, and those with a predilection for gravitational units have no right to relegate the poundal to a subsidiary position as an auxiliary unit only (see later, Clause 20, B.A. *Report*, 1907). The poundal is not an educational convenience —it is a unit of force philosophically conceived and quite independent of the gravitational attraction of the earth. It does not vary from place to place, but is a fixed unit which may be used quite naturally and universally in all dynamical problems.

The system of units used by engineers is not happily conceived. A horse-power is 33,000 foot-pounds per minute. This necessarily involves the pound weight as the unit of force, and the name for the corresponding unit of mass has not come into common use. Occasionally, the engineer borrows a unit of mass from the F.P.S. system as a matter of popular convenience, and this leads to chaotic units. If the density of water is quoted as 62·5 pounds per cubic foot on the back of a slide rule, this is pounds mass: otherwise it would be dimensionally incorrect. The pound mass properly belongs to the F.P.S. system. Again, the British thermal unit involves the pound mass. In working out the mechanical equivalent of heat from Joule's original experiment, the heat is reckoned in British thermal units involving pounds mass, and the work is in foot-pounds involving pounds weight. If the engineer must use these units (and it is not expected that he will change them) he should recognize that he has borrowed the pound mass for the time being from the F.P.S. system, but on no account should such units used together be regarded as an alternative " system " of units.

In the Preface to *Hydrostatics and Elementary Hydrokinetics*, by George M. Minchin, the following passage occurs:—" Some long controversies have recently taken place between the representatives of the practical engineers, on the one hand, and those who may be termed the scientific reformers, on the other, with regard to the way in which we should speak of the common gravitation measure of force. The former speak of ' a force of one pound ', for example, while the latter prefer ' a force of one pound weight '. As in the last edition of my Statics, I have everywhere throughout the present work adopted the latter mode of speaking, because I regard the distinction between *mass* and *weight* of a body as absolutely fundamental—the former being an essential and invariable, while the latter is a purely contingent and variable, attribute of the body. Indeed, were it not for the fact that the Earth is nearly spherically symmetrical as regards shape and density about its centre, it seems scarcely possible that men, communicating with each other over long distances, could ever have adopted the former mode of speaking " 1

Those who receive elementary instruction in mechanics will ultimately be trained as mathematicians, astronomers, physicists, engineers of all kinds, civil, mechanical, electrical, aeronautical, marine, etc. All these students receive their early training together, and their early studies should lead quite naturally and without difficulty to any more advanced form of training. Those teachers whose training has been specialized and who attach great weight to gravitational units should co-operate to the full in an attempt to ease the path of the beginner without interfering with the rigour of his training or narrowing his outlook on this important question of units.

The three systems of units which should be taught to all students from the very beginning are the C.G.S., the F.P.S. and the gravitational systems. A student should work problems in all these units. There will be some difference of opinion with regard to the inclusion of the F.P.S. system. This difference of opinion is a sure and certain indication that units are not being taught in the right way. If the treatment of units is satisfactory, the inclusion of F.P.S. units will help rather than hinder the student in getting a clear grasp of his problems. A little extra time must be devoted to the matter, but this is time well spent. Any student who can use the C.G.S. system has no difficulty whatever with the F.P.S. system, and conversely. Still further, if the gravitational unit of mass is given a name in the same way as the other units, any student who knows how to use one system has no difficulty with any other system. This is why uniformity of treatment is so essential. The treatment of any problem and the symbols and formulae used should be the same whichever set of units is used. It is most undesirable to develop the subject by any method which requires modification in passing from one system of units to another or which requires modification in any way at a more advanced stage of study. Moreover, a student should use all systems of units with equal facility and in the same way.

§ 3. UNITS OF MASS

The standard pound and the standard kilogram are the legal units of mass (see Quotation 59). A gram is one thousandth of a kilogram. 1 cubic centimetre of water at 4° c. weighs approximately 1 gram, in fact 1·0000 cubic centimetre of water at 4° c. and at a pressure of 1 standard atmosphere weighs 1·0000 gram. Reference to the temperature of maximum density of water should be avoided. For greater accuracy of expression, the isotopic composition of water should be specified.

A litre used to be regarded as a cubic decimetre or one thousand cubic centimetres. Why change this definition? The graduation marks on a burette or a pipette are not visibly altered from graduation marks in cubic centimetres, and, in any case for accurate work, these instruments will be calibrated by weighing. It would seem desirable to return to the original definition of the litre and for measuring vessels to be graduated in c.c. instead of ml. As things now are, the instrument maker appears to be the only one concerned with the millilitre. Measuring vessels are graduated in millilitres, and the user calls them cubic centimetres.

§ 4. THE CONCEPT OF MASS

The notion of mass, apart from its exact quantitative definition, is not a difficult one. The inertia of objects and the quantity of matter in food, drink and raw materials are well known to everyone and matters of everyday experience. Rough estimates of mass are commonplace.

E. Mach, in *The Science of Mechanics*, states:—"With regard to the concept of ' mass ', it is to be observed that the formulation of Newton, which defines mass to be the quantity of matter in a body as measured by the product of its volume and density, is unfortunate. As we can only define density as the mass of unit volume, the circle is manifest. Newton felt distinctly that in every body there was inherent a property whereby the amount of its motion was determined, and perceived that this must be different from weight. He called it, as we still do, mass; but he did not succeed in correctly stating this perception." 2

Consider a heavy object (a 56-lb. weight or large bucket of shot or sand) suspended by a long and strong cord from a rigid and high support. The weight of the object is balanced by the tension in the cord. Horizontal displacement or rotation of this massive object is resisted. Motion of translation or rotation cannot readily be reduced or increased. The law of inertia relates to this universal property of matter of resisting change in its state of rest or of uniform motion. This universal property of matter and a constant property for any particular object is called inertia and is measured as mass.

Newton's first law of motion provides a definition of mass and a definition of force.

§ 5. DEFINITION OF MASS

Science of Mechanics, by E. Mach, 1907.

" In the first place we do not find the expression ' quantity of matter ' adapted to explain and elucidate the concept of mass, since that expression itself is not possessed of the requisite clearness." 3

" If, however, mechanical experiences clearly and indubitably point to the existence in bodies of a special and distinct property determinative of *accelerations*, nothing stands in the way of our arbitrarily establishing the following definition:—

" All those bodies are bodies of equal mass, which, mutually acting on each other, produce in each other equal and opposite accelerations." 4

This definition of equal masses is capable of illustration with a ballistic pendulum.

" No *logical* necessity exists whatsoever, that two masses that are equal to a third mass should also be equal to each other." *Experience* shows that they are. 5

" All uneasiness will vanish when once we have made clear to ourselves that in the concept of mass no theory of any kind whatever is contained, but simply a fact of experience." 6

Newton's second law of motion leads to the relation

$$P = kma,$$

where k is a constant which is equal to unity if unit force produces unit acceleration in unit mass. In this case $P = ma$, a relation which is true for any system of units. This relation defines the unit of force if the unit of mass has already been defined as in the C.G.S. and F.P.S. systems. In the gravitational system, the unit of force is defined first and the relation $P = ma$ then defines the unit of mass. All the fundamental units and most of the derived units have a name in all systems of units. Surely some name should be given to the unit of mass in the gravitational system. The name *slug* has been suggested and might well be adopted.

§ 6. THE GRAVITATIONAL UNIT OF MASS

It seems urgently desirable for those who use the gravitational system of units to use a set of consistent units. If the pound weight is the engineers' unit of force, the pound mass should *not* be regarded as the engineers unit of mass. The user of gravitational units sometimes borrows the pound mass from the F.P.S. system as a matter of popular convenience and not as a matter of philosophical virtue or scientific rigour. The use of the pound as a unit of mass as well as a unit of force cannot fail to cause confusion and difficulty, particularly in the case of beginners. Such chaotic use of units should be recognized as chaotic and unscientific, and as such should be avoided. No conjoint use of chaotic units should be called a system of units.

Quotations 44 to 51 are from the writings of well-known engineers and professors of engineering. These quotations indicate that the engineer is fully alive to the importance of mass in engineering discussions and in dynamical equations. Professor Perry says " we have no name for unit of mass ", and attempts to justify this by adding " the engineer never has to speak of the inertia of a body by itself". (Quotation 46 indicates that the term " inertia " is here used as synonymous with " mass ".) This last remark is quite inconsistent with Perry's treatment of the inertia of moving parts of an engine. In Perry's *Steam Engine* (1899), Chapter XXIX, many references to mass occur; for example:—

" If it were possible to imagine the effect of the mass of the connecting rod to be the same as that of two masses at its ends, it would be easy to balance engines." 7

" A mass m whose centre of gravity is at a distance r from the axis: it is between the wheels at the lateral distances l_1 and l_2. What masses of the wheels will balance it? " 8

" There will be a mass of $\frac{276}{32\cdot2} \times \frac{7}{15}$ or 4·00 moving with acceleration $\ddot{s}$ and a mass $\frac{276}{32\cdot2} \times \frac{8}{15}$ or 4·56 on the crank-pin." 9

In general, users of gravitational units do not give the gravitational unit of mass a name. In the preface to Professor Worthington's *Dynamics of Rotation* (1902), he says:—

" In the interests of clear teaching . . . . I have ventured to give the name of a slug to the British engineers' Unit of Mass." 10

Mr. Awbery tells me that the slug is used in the hundreds of reports and memoranda of the Aeronautical Research Committee and many other aeronautical publications. Why has not this name been adopted universally? Professor Worthington was one of the Committee responsible for the B.A. *Report*, 1907. This Report makes no reference to a matter of fundamental urgency and importance—a name for the gravitational unit of mass. Quotations 14 and 15 indicate the difficulty which the engineer has in expressing a quantity such as moment of inertia in the absence of any name for the unit of mass which is involved.

§ 7. MOMENT OF INERTIA

Quotations 11 and 12 indicate the unfortunate consequences of using pound weight and pound mass (the same name for units of different dimensions). All formulae must be duplicated and still there is confusion, due to chaotic units.

W. J. M. Rankine, *Applied Mechanics*, 1882.

" The moment of inertia of an indefinitely small body, or ' physical point ' relatively to a given axis, is the product of the mass of the body, or of some quantity proportional to the mass, such as the weight, into the square of its perpendicular distance from the axis: thus, in the following equation,

$$I/g = mr^2 = Wr^2/g.$$

r is the perpendicular distance of the mass m, whose weight is W, from a given axis; and the moment of inertia, according to the unit employed, is either I or I/g; the former when the unit is the moment of inertia of an unit of *weight* at the end of an arm whose length is unity, and the latter when the unit is the moment of inertia of an unit of *mass* at the end of the same arm. For the purposes of applied mechanics the former is the more convenient unit, ... " . . 11

" In British measures, moment of inertia will be expressed by the product of a certain number of pounds avoirdupois into the square of a certain number of feet.

$$I = \Sigma \,.\, Wr^2, \quad E = a^2I/2g,$$

where a = angular velocity, E = energy of rotation about an axis through the centre of gravity." 12

Louis Toft and A. T. J. Kersey, *Theory of Machines*, 1939.

“ $I=\Sigma . mr^2$,’ $C=Ia$ ” 13

“ I is known as *moment of inertia*, and if m is in engineers’ units, and r in feet, I is measured in mass (feet)^2 units.” 14

In the worked examples, moments of inertia are given in engineers’ units, the mass being reckoned in engineers’ units.

W. E. Dalby, *Steam Power*, 1915.

An answer to a numerical example on a fly-wheel is stated as “6·34 W/g. ft^2 units.” 15

Quotations 14 and 15 indicate that a name for the engineers’ unit of mass is required. A mass could then be stated in slugs and a moment of inertia in slug-feet^2

The following formulae apply to any consistent system of units:—

Any theorem in linear motion may be applied to angular motion by making the following substitutions:—

Moment of Inertia I for Mass m.
Torque or Couple G for Force F.
Angular displacement and its derivatives θ, ω, α for linear displacement and its derivatives s, v, a.

Thus $G=I\alpha$ is parallel with $F=ma$,
K.E. $=\frac{1}{2}I\omega^2$,, ,, K.E. $=\frac{1}{2}mv^2$.
Work $=G\theta$,, ,, Work $=Fs$.

If chaotic units and formulae are rejected, one formula is applicable in any system of units.

§ 8. DENSITY

Quotation 59 is a model of verbal and legal accuracy. Unfortunately, what is here called the Weight is only the Weight where nobody ever weighs. The operation of weighing described in Quotation 59 always yields the same result for the same body, no matter where it is carried out. The result of the operation of weighing on a beam balance is by general scientific consent called the mass of the body. The word mass should be on the tip of the tongue of the beginner all the time, his note-book entries should record mass, and the word weight should only be used when the gravitational attraction of the earth is involved.

The only definition of density which is numerically and dimensionally true is mass of unit volume, $[D] = [ML^{-3}]$. I do not favour any alternative. A density in pounds per cubic foot is in F.P.S. units; density is not pounds weight per cubic foot. If an engineer expresses a density in pounds per cubic foot, he has borrowed the F.P.S. system as a matter of popular convenience, but pounds weight per cubic foot is not correct.

The engineer uses the pound weight as the unit of force, but on occasion he uses the pound as the unit of mass. As remarked above, if the density of water is stated to be 62·5 pounds per cubic foot on the back of a slide rule it is not certain whether the pound weight is indicated or the pound mass, nor, for popular practical purposes, is it necessary to know. In this case, slugs per cubic foot would not be so practically useful, but it is only the expression of density which is numerically and dimensionally true for use in dynamical formulae and equations, unless the F.P.S. system is used.

For example, quoting from Kermode's *Mechanics of Flight*:—

"Our formula becomes $R = K\rho AV^2$, R being in lb. force, ρ in slugs per cub. ft. (approx. 0·0024 slugs per cubic foot under standard conditions), A (sq. ft.), V (ft./sec.) and K (no units)." 16

Whatever apparent confusion of units and dimensions there may be in engineering calculations, this is not due to lack of care or accuracy on the part of engineers. Writers on engineering subjects usually state their results with great care and accuracy; for example, statements 14 and 15 look clumsy, but there are no better ways of expressing the result in the absence of a name for the engineers' unit of mass.

§ 9. SPECIFIC GRAVITY

Specific gravity is relative density—the ratio of the density of a substance to the density of a standard substance (usually water at 4° c. or air or hydrogen at S.T.P.). When the chemist says that the density of a gas is half its molecular weight, he means specific gravity and not density. Evidently specific gravity is equal to the mass of the substance divided by the mass of an equal volume of standard substance or the volume of a certain mass of the standard substance divided by the volume of an equal mass of the substance.

How is the Principle of Archimedes to be stated? The principle deals with forces, not masses. It is the apparent loss of weight which is equal to the weight of liquid or gas displaced. But measurements made with a beam balance are not the weights, but the masses which are proportional to these weights. The use of the term gram weight should be avoided.

Example: Using a beam balance, a solid weighs 100 grams in air, 80 grams in water, and 75 grams in a liquid. Find the density of the solid and of the liquid.

Mass of solid in air, 100 grams.
Mass of solid in water, 80 grams.
Mass of solid's volume of water, 20 grams.
Mass of solid's volume of liquid, 25 grams.
Specific gravity of solid = 100 grams/20 grams = 5.
Specific gravity of liquid = 25 grams/20 grams = 1·25.
Density of solid = 5 grams per c.c. or 5 × 62·5 lbs./c.ft.
Density of liquid = 1·25 grams per c.c. or 1·25 × 62·5 lbs./c.ft.

In this example, if the measurements are in pounds the wording and calculation remain the same if pound is substituted for gram.

If a spring balance is used, weight is substituted for mass.

§ 10. PRESSURE

Pressure (P) is force (F) per unit area (A).

Pressure may be stated in dynes per square centimetre, or in poundals per square foot or in pounds weight per square foot.

Pressure in grams weight per square centimetre should be avoided. Head of water is more satisfactory. Pressure in pounds per square foot obviously means pounds weight per square foot, and is shorter for marking on a pressure gauge.

§ 11. THE CHAOTIC GRAM WEIGHT

It will be generally conceded that much confusion has arisen with regard to units and formulae owing to the use of the pound weight and the pound mass in the same problem or formula. This chaotic use of units is now so firmly established in connection with certain types of problem that the difficulty cannot now be entirely eradicated. It is, however, desirable for the student to recognize chaotic units whenever they appear, and to convert them into rational units at once. It is important to change the chaotic unit first and to manipulate the rationalized unit later. Special chaotic formulae for chaotic units naturally lead to utter chaos.

It seems obvious that the introduction of chaotic units is undesirable, and particularly is this the case when the gram weight is used as a unit of force. The dyne is the C.G.S. unit of force and the gram is the C.G.S. unit of mass. The gram weight is a chaotic unit, unwanted and entirely unnecessary. Past experience, quite apart from other important considerations, should have prevented the introduction of such a dangerous source of confusion. The gram weight should be rigorously rejected and carefully avoided.

Similarly, a pressure in grams weight per square centimetre is not rationally expressed. In C.G.S. units a pressure is in dynes per square centimetre. To express forces in grams weight or pressures in grams weight per square centimetre is to acknowledge defeat in the application of a rational system of units (C.G.S. or F.P.S.). Already there are at least four different styles of working in pound units. Are there to be four more in gram units? Will the dyne mass be invoked at a later stage? The dyne mass is no more and no less chaotic than the gram weight.

The use of the gram weight as a unit of force is far worse than the use of the pound mass occasionally by those who more often use gravitational units. The gram weight is either entirely unnecessary and chaotic or it is used in preference to the dyne. The teacher of mechanics who uses the formula $P/W=a/g$ seeks refuge in the gram weight on those very rare occasions when problems are considered which should be worked in C.G.S. units. He is aware that utter confusion prevails when this formula is used to obtain a result in dynes, partly because such results are seldom contemplated. The introduction of the gram weight as a unit of force is a confession of failure to use C.G.S. units. Most of those educated in this way will be quite incapable of studying quantitative magnetism and electricity until the whole of the past pernicious training has been eradicated. It will then be necessary to start afresh with units and finally to deal with unit pole, etc. Even so, the position of the student is always precarious. There may be regression at any time and consequent confusion of ideas.

To reaffirm a matter of the highest importance. There is no reason whatever to use the gram weight as a unit of force or to express pressures in grams weight per square centimetre. There is every reason to discountenance these spurious units. They are used solely by those who employ methods and formulae only applicable to gravitational units and who seldom work a problem in C.G.S. units except under compulsion. Such half-developed teaching should be penalized. Every student of mechanics should be able to work problems in C.G.S. units—forces should be in dynes and pressures in dynes per square centimetre. No other units of force or pressure are necessary or should be tolerated. It should not be necessary for a teacher of magnetism and electricity to explain the meaning of the term *dyne* to those who have been studying dynamics. The matter becomes serious if the science teacher finds it impossible for the past training of the student in dynamics to adapt itself to scientific uses.

Lancelot Hogben, *Science for the Citizen*, 1937.

" Force = Mass × Acceleration.

" 1 poundal = force required to impart accel. of 1 ft. per sec^2 to mass of 1 lb.

" 1 dyne = force required to impart accel. of 1 cm. per sec^2 to mass of 1 gm." 17

" In the international system, the unit of force is one *dyne*." 18

§ 12. NEWTON'S LAWS OF MOTION

Professor Thomas Case, in *Lectures on the Method of Science*, refers to Newton's laws as follows:—

" Newton begins with the three axioms or laws of motion: the law of inertia of rest and motion, the law of proportion of force to effect, and the law of reciprocity, or equal action and reaction. He gives inductive evidence of these three laws. Though we know of no body altogether unimpeded, yet there are in our experience many instances of bodies moving straight forward the more freely the less they are impeded, which enable us to deduce the law of inertia in the form that every body perseveres in its state of rest or motion *so far* as it is not affected by an impressed force. Secondly, the proportion of force to effect is obvious to experience. The third law had been already deduced for the force of impact by Sir Christopher Wren's experiments before the Royal Society, which proved that, if one body impinge upon another, and by its force change the motion of the other, that first body also, because of the equality of the mutual pressure, will undergo an equal change in its own motion towards the contrary part. Newton repeats Wren's experiments, and, having added fresh instances from magnetism and other attractions, extended the law from impact to all forces, and induced it in the universal form, action and reaction are always equal and opposite. Thus did Newton finally establish the main principles of dynamical mechanics by the empirical method of Bacon. But he went beyond induction, which never could have discovered the imperceptible link between the sun and its planets. The sequel is deduction; and even under the head of *Axioms* he immediately uses the third law, as induced from impact, to deduce that the whole momentum of bodies acting reciprocally on one another, together with the state of their centre of gravity, is conserved. So true is it that induction leading to deduction is the most fruitful method of natural science." . . . 19

Professor F. Gotch, in *Lectures on the Method of Science*, says:—

" Newton's enunciation of the law of gravitation is so familiar that I might merely mention it, but I desire to do more than this, because I do not find myself in complete accord with the view already put before you by my colleague, Professor Case, who has selected this as a typical example of logical deduction. For my own part I regard Newton's statement that every particle of matter in the universe exercises an attractive force on every other particle of matter, which varies inversely as the square of the distance between them, as an induction, the conception of a creative mind gifted with imagination. Professor Case himself stated that Newton passed from terrestrial to celestial mechanics. In the language of Tyndall, this ' passage from a falling apple to a falling moon ' was a stupendous leap of the imagination, for his enunciated law applies in conception to the universe, thus extending into boundless space and persisting through endless time." 20

Space does not permit any further general discussion of Newton's laws, but admirable expositions are to be found in Cox's *Mechanics* and Mach's *Science of Mechanics*. Quoting from Mach:—

" Newton's sense of *what* fundamental concepts and principles were required in mechanics was admirable. The *form* of his enunciations, however, as we shall later indicate in detail, leaves much to be desired. But we have no right to underrate on this account the magnitude of his achievements; for the difficulties he had to conquer were of a formidable kind, and he shunned them less than any other investigator." 21

§ 13. THE FUNDAMENTAL DYNAMICAL RELATION

Newton's second law indicates that the ratio of the force (F) acting on a body to the acceleration which that force gives to the body (a) is constant.

The ratio F/a is the mass of the body.

Hence $F = ma$, where m is the mass of the body.

Thus mass is the coefficient of acceleration in the fundamental dynamical relation $F = ma$ (see Quotation 74).

Since the weight (W) of a body acting on its mass m produces acceleration g of the body, $m = W/g$.

In engineers' units:

$$\text{Mass in slugs} = \frac{\text{weight in pounds weight}}{32{\cdot}2 \text{ ft./sec}^2}.$$

For example, 64·4 pounds mass = 2·00 slugs.

The relation $F = ma$ indicates that unit force produces unit acceleration in unit mass. Thus,

1 dyne produces 1 cm./sec^2 acceleration in mass 1 gram.
1 poundal produces 1 ft./sec^2 acceleration in mass 1 pound.
1 pound weight produces 1 ft./sec^2 acceleration in mass 1 slug.

The first statement defines the dyne, the second defines the poundal, and the third defines the slug.

Thus, if the corresponding units of work be added, the following table of units is obtained:—

System	Length	Mass	Time	Force	Work
C.G.S.	CENTIMETRE	GRAM	SECOND	Dyne	Erg
F.P.S.	FOOT	POUND	SECOND	Poundal	Foot-poundal
Gravitational	FOOT	Slug	SECOND	POUND WEIGHT	Foot-pound

Fundamental units are printed in small capitals.

The formula $F = ma$ is in the simplest possible form for direct use and concrete comprehension. Any attempt to modify this relation in order to introduce weight (W) into it instead of mass naturally leads to confusion of units and unnatural thinking. $F = ma$ is the only relation which is appropriate to problems dealing with force and acceleration.

A student may be called upon to work a problem stated in chaotic units. This is no excuse for introducing chaos into the fundamental relation. The procedure should be to remove the chaotic units and then apply the formula $F = ma$. For example: What uniform force produces an acceleration of 10 ft./sec^2 when acting on a body of mass 64·4 pounds?

This question is set in F.P.S. units. The natural solution is

$$m = 64{\cdot}4 \text{ pounds}, \ a = 10 \text{ ft./sec}^2, \ F = ma = 644 \text{ poundals.}$$

If the answer is expected in pounds weight the question should be worded: What force in pounds weight produces an acceleration of 10 ft./sec^2 when acting on a body of mass 64·4 pounds ($g = 32{\cdot}2$ ft./sec^2).

The properly trained student will immediately sense chaotic units. The result is required in gravitational units. Hence all quantities must be expressed in gravitational units before substitution in the formula $F = ma$:

$$m = 64{\cdot}4/32{\cdot}2 \text{ slugs} = 2{\cdot}00 \text{ slugs}, \ a = 10 \text{ ft./sec}^2$$
$$F = ma = 2{\cdot}00 \times 10 \text{ pounds weight} = 20 \text{ pounds weight.}$$

If, in the previous question, the weight of the body is 64·4 pounds weight, the units are not chaotic. Since $W=mg$, .

$$m=64{\cdot}4/32{\cdot}2=2{\cdot}00 \text{ slugs.}$$
$$F=ma=2{\cdot}00\times 10=20 \text{ pounds weight.}$$

It is emphasized that chaotic units should be corrected before substituting in the formula $F=ma$. It is *not* advocated that chaotic units should induce a student to use a chaotic formula.

§ 14. THE B.A. REPORT OF 1907

" *The Teaching of Elementary Mechanics.*—Report of the Committee, consisting of Professor Horace Lamb (Chairman), Professor J. Perry (Secretary), Mr. C. Vernon Boys, Professors Chrystal, Ewing, G. A. Gibson, and Greenhill, Principal Griffiths, Professor Henrici, Dr. E. W. Hobson, Mr. C. S. Jackson, Sir Oliver Lodge, Professors Love, Minchin, Schuster, and A. M. Worthington, and Mr. A. W. Siddons, appointed for the Consideration of the Teaching of Elementary Mechanics, and the Improvement which might be effected in such Teaching.

" Clause 11. The impression that the weight of a body is in reality a single force acting at its centre of gravity should be guarded against.

" Clause 19. It is convenient to treat elementary problems on the accelerations produced by forces by simple proportion,

$$\frac{\text{force acting}}{\text{weight}}=\frac{\text{acceleration produced}}{g},$$

using the fact that a body's own weight produces acceleration g; and it is convenient to postpone the consideration of mass or inertia until such problems have been discussed.

" Clause 20. Students ought to know the meaning of ' absolute measure '; that is, they should be able to interpret all fundamental equations such as $E=\frac{1}{2}mv^2$, or $F=mdv/dt$, in any consistent system of units whatever. They should learn that for certain purposes the C.G.S. system of units is convenient; and for certain other purposes the units employed by British engineers may be convenient; they should not be dominated by any system, but able to use them all. The *poundal* and other such educational conveniences should be used as auxiliary units only, final results being expressed in units to which practical people are accustomed, so as to be generally intelligible."

Clauses 19 and 20 are extraordinary in that so many members of the Committee responsible for this Report have never indicated agreement with these clauses in their writings (see Appendix 1).

Clause 19 advocates the use of the relation $F/W=a/g$. Clause 19 also suggests that the idea of mass should be shelved until problems on $F/W=a/g$ have been discussed.

I venture to think that Clause 19 has done more mischief in the teaching of mechanics than the aggregate of all other ill-conceived suggestions that have ever been made in the teaching of this subject. The effect of these suggestions has led to the appearance of school textbooks of mechanics which do not mention the term mass at all, and which ignore C.G.S. units entirely. In others, the total space devoted to mass and C.G.S. units is about three consecutive pages, including the examples. What is so obvious in every case in which the formula used is $F/W=a/g$ is that it is quite unsuitable for dealing with any system of units except the gravitational system. Again, there are textbooks of science in which, after a whole chapter has been devoted to inertia and the formula $F=ma$ has been developed quite happily, there comes a sudden urge to pander to the recommendations of this *Report*, and an unnatural twist is given to the treatment in order to mention the formula $F/W=a/g$. Some teachers of mechanics use two sets of formulae throughout, one for C.G.S. units and the other for gravitational units. The cast-iron recipe is: if the body weighs m grams use $F=ma$, and if the body weighs W pounds weight use $F/W=a/g$.

Clause 20 recommends that a student should be able to interpret $E=\frac{1}{2}mv^2$ and $F=mdv/dt$ in any consistent system of units whatever. $F/W=a/g$ is the worst possible preparation for this purpose, and leads to confusion when C.G.S. or F.P.S. units are used. How is this formula to be applied to C.G.S. or F.P.S. units?

Consider the simple problem: "What force acting on a body weighing 10 grams produces an acceleration of the body of 20 cm./sec^2?

$W=mg=10\times 981$ dynes (using the formula $F=ma$—or is there any alternative to the use of this formula?).

$$a=20 \text{ cm./sec}^2,\quad \frac{F}{10\times 981}=\frac{20}{981},$$

giving $F=200$ dynes.

Can anything be more fatuous?

Alternatively, $W=10$ grams weight, $a=20$ cm./sec^2, $F/10=20/981$, giving $F=200/981$ grams weight.

This way of working introduces an unwanted unit of weight into the C.G.S. system, which seems worse than the previous method.

The formula suggested in Clause 19 is not consistent with the recommendation to use any system of units contained in Clause 20.

Clause 20 continues: "The poundal and other such educational conveniences . . . . generally intelligible."

This statement has already been considered in part. The poundal has been used by many mathematicians and physicists for philosophical rigour and not for educational convenience. Moreover, the statement under examination is inconsistent with the recommendation made earlier that "students . . . . should be able to interpret all fundamental equations . . . . in any consistent system of units whatever."

A student properly trained in the use of all units will express his result in the unit most appropriate to the problem or, maybe, in more than one way. The practical people mentioned in the *Report* are not defined, and their predilection for any particular system of units is somewhat indefinite, but presumably the same as the writers of the *Report*.

In view of Clause 11, what mental picture is a student supposed to conjure up when working a simple problem on force involving the application of $F/W = a/g$? Or, is the formula to be applied without forming any mental picture of the problem? What diagram can a student draw to help him to crystallize his ideas and data?

The formula $F = ma$ involves no ratios at all. A spring balance or a cord passing over a pulley to a weight F pulls the body horizontally along a smooth surface, and its displacements after equal intervals of time show ever increasing amounts. A student can visualize motion according to $F = ma$, but how can he picture the process $F/W = a/g$? How can a student visualize two ratios simultaneously?

It would be very interesting to know more details of the discussion which preceded the publication of this *Report*, and particularly that relating to Clauses 19 and 20. What were the views of various members of the Committee and to what extent was opinion divided with regard to Clauses 19 and 20? Various quotations in Appendix 1 indicate the views of some of the members of the Committee as expressed in their writings.

I have been somewhat critical of Clauses 19 and 20 of the *Report*. which relate to the teaching of " Mass ". I should, however, support most of the other recommendations.

§ 15. THE FORMULA $F/W = a/g$

This formula is undesirable for the following reasons:—

1. It is not suitable for use with C.G.S. or F.P.S. units.

2. The formula introduces the weight of the body into problems in which the weight of the body is not directly involved.

3. The formula avoids all reference to mass.

4. The formula separates quantities which should naturally come together. $F = Wa/g$ is better, but this is $F = ma$.

5. This formula and all other formulae in the dynamics of translation and the dynamics of rotation must be duplicated unless the student is confined to gravitational units.

6. Those who use this formula commonly presuppose a late development, or the complete elimination of, the idea of mass.

7. It is impossible to visualize the simplest force problem in terms of this formula.

8. The formula lacks experimental and historical background.

The only formula readily applicable to all systems of units is the simplest possible formula, $F=ma$.

It is of interest to trace the origin of the formula $F/W=a/g$ and to ascertain the consequences of its employment. The first extensive use of this formula appears to have been in the Ordnance College, Woolwich, by Professor Sir George Greenhill and the Instructors at the Royal Military Academy, Woolwich, C. S. Jackson and W. M. Roberts. Teachers of mathematics in public schools preparing candidates for the Army Entrance Examination found this formula quite sufficient for dealing with all the problems set in the various grades of mathematical paper because problems were almost invariably set in gravitational units only.

Greenhill's *Notes on Dynamics* was published by H.M. Stationery Office. These *Notes* mark a distinct advance in the teaching of dynamics as applied to practical problems. The book teems with examples which are in a class by themselves, and the treatment is brilliant throughout. Quoting from this book:—

" We begin Dynamics by considering motion in a straight line, as of a train on a straight railway, a shot in the bore of a gun, the piston of a steam engine, or water in a straight uniform pipe. This may be called the Dynamics of one dimension, and much use is made of it in the dynamical interpretation of facts familiar in everyday experience." 22

It is most remarkable that this book, so brilliantly original and so versatile and prolific in practical illustrations of every kind, should develop the subject of Dynamics without one single allusion to the term " mass ". It is also to be noted that the book contains no force or work problems in C.G.S. units.

Quoting from the section on *Moment of Inertia*:—

" Definition. The moment of inertia (M.I.) of a body about an axis is the sum of the product of the weight of each particle of the body and the square of the distance from the axis.

" Thus if m lb. is the weight of a particle at a distance r ft. from the axis, the M.I. of the particle is mr^2, in lb. ft^2, and of the whole body is Σmr^2 (lb. ft^2), . . . " 23

The whole book shows a most careful and studious avoidance of the idea of mass. (See Quotation 59).

The *Notes on Dynamics* were followed by *A First Dynamics*, by C. S. Jackson and W. M. Roberts. This book contains an excellent collection of examples of all kinds, but hardly any at all on C.G.S. units. Quoting from the preface to the book:—

" There is still, perhaps, a danger that English text-books on Elementary Dynamics may become sharply divided into two classes, to which the dyslogistic epithets of ' academic ' and ' practical ' have been applied, the one for the public school-boy and the candidate for University Scholarships, the other for the technical student and the engineer. We regret the existence of this distinction, and have tried to be neither ' academic ' nor ' practical '." 24

" On account of its difficulty we have postponed the introduction of the notion of Mass until a later stage than has been usual." 25

Mass is not referred to until Chapter X on Newton and the Idea of Mass.

The following are quotations from Chapter X :—

" The student has realised that the notion of mass is not required for the dynamics of the phenomena of ordinary life on the earth's surface." . . . 26

As has already been pointed out, mass is more fundamental to ordinary experience than weight.

" For scientific purposes the centimetre, gramme, second (C.G.S.) system is invariably used. Unfortunately, this point of view has not always been adopted." 27

" Much confusion has been caused by rejecting the scientific unit of mass, the gramme." 28

There are 105 Miscellaneous Examples at the end of the book. Not one of these deals with C.G.S. units.

The next stage in the argument consists in deducing the formula $P=kmf$ from the formula $P/W=f/g$ (see pp. 208, 209).

" Summary.—The same fundamental dynamical relation between force and acceleration has been expressed in different ways (for British units):

" (1) $P/W=f/g$. P and W in the same of any units of force, pounds, tons, etc.

" (2) $P=1/32\,mf$. m in pounds mass, P in pounds weight.

" (3) $P=mf$. P in pounds, m in slugs.

" (4) $P=mf$. P in poundals, m in pounds mass.

" It is worth noticing that, if for any purpose we wish to pass from (1) to *any* system of units involving mass in which the equation $P=mf$ is true, we have only to replace W by mg.

" For this reason we have no hesitation whatever in advising the student to adopt (1) as his standard dynamical equation for British units." 29

It will be observed that C.G.S. units are ignored. Again, formula (2) is chaotic—it should be rejected. Formula (1) rearranged is $P=Wf/g=mf$.

Thus, $P=mf$ is a relation which applies to any consistent system of units without any modification whatever.

Other text-books of dynamics have now appeared using the formula $P/W=a/g$. All these books reveal the same characteristics :—

1. They are written by mathematicians who have had little scientific or engineering training.

2. They either ignore C.G.S. units completely or the treatment is confined to a short paragraph and the examples are all together and entirely unrepresentative.
3. The idea of mass is either ignored completely or dealt with unsatisfactorily.
4. The treatment bears no relation whatever to the previous or future scientific training of the student.

Science text-books on *Mechanics* and *Hydrostatics* seldom emphasize the distinction between mass and weight at an early stage, or attend to the rigour of the treatment in problems on density and specific gravity. At a later stage, the formulae $F = ma$ and $F/W = a/g$ are both developed in order that the book may appeal to a wider market. This duplication of formulae and treatment is to be deplored.

§ 16. EXAMINATIONS

Examination syllabuses prescribe the subject matter to be taught and often indicate the type of treatment required. The careful setting of examination questions is of great importance and assistance in developing the subject in the right way. A student taught on historical lines should not expect questions on the history of the subject. Similarly, the student trained by the experimental method must not expect direct questions on dynamical experiments. He should be tested on the principles of the subject, and any excellence in the methods of teaching will be reflected in his grasp of those principles.

The early distinction between mass and weight, and the careful and correct use of these terms, should be emphasized in all examination syllabuses not only in physics but in mechanics (additional mathematics, applied mathematics, etc.).

Questions should be set in all systems of units, and the candidate should be expected to use all systems with equal facility, not only in " Science " but also in " Mechanics." This is extremely important, especially in the mechanics papers, where there is a tendency to set gravitational units only.

Some uniform system of notation and formulae is desirable.

Beam-balance weighings should be recorded as mass, and spring-balance readings as weight. The terms *mass* and *weight* should be used carefully and correctly in the working out of problems on density and specific gravity.

Forces in grams weight and pressures in grams weight per square centimetre should be rejected. The gram weight is unwanted and unnecessary, just as the dyne mass would be equally unwanted and unnecessary. These are not C.G.S. units, and inevitably lead to confusion. A pressure is expressed in dynes per square centimetre or as head of liquid.

Mass and weight should enter into formulae and equations naturally, so as to lead to dimensionally correct results.

Values of g should be given at the head of the paper and not at the end of the question. For approximate or rapid calculation $g = 32$ ft./sec^2 or, very roughly, 1000 cm./sec^2, which is approximately 33 ft./sec^2

APPENDIX 1

Attention is directed to quotations from the writings of members of the Committee responsible for the B.A. *Report* of 1907. It is probable that no member of this Committee agreed entirely with the whole of Clauses 19 and 20. Most of the members probably agreed with the first part of Clause 20 and would not favour Clause 19. The teachers of *Mathematics Applied to Military Engineering* would probably agree with Clauses 19 and 20 except the first part of Clause 20.

QUOTATIONS FROM ELEMENTARY TEXT-BOOKS

Oliver J. Lodge, *Elementary Mechanics, including Hydrostatics and Pneumatics* (1897.)

" . . . . it is hoped that students who use this manual will be able to master the elements of the science in such a way that they may rise from it to more advanced treatises, not only without having anything to unlearn, but with a very sound knowledge of principles." 30

" Experiments in Mechanics have a subordinate though most useful part in illustrating and emphasising the facts, but the author has no faith in making the establishment of principles depend on special experiments." 31

" The book, ... is intended to be not only an easy introduction to the subject, but, as far as it goes, a philosophical work." 32

" To recapitulate, then, mass means quantity of matter, and is measured by inertia." 33

The relation $F=ma$ is used. Examples illustrating the use of C.G.S., F.P.S., and gravitational units are given throughout the book.

W. D. Eggar, *Mechanics* (1905).

" The writer has found that Kinetics as well as Statics may be helped immensely by experimental work, and that the trouble involved in the preparation of lecture-experiments is amply repaid by the clearer notions gained by students on such points as acceleration, energy, and momentum." 34

" . . . . apart from the historical interest, Galileo's methods were so simple and accurate, his mathematics so sound and his arguments so lucid, that they ought to be more familiar. Experiments which resulted in the discovery of the laws of motion ought to be worth repeating." 35

" An experimental verification of Newton's Second Law has been given. To justify this rash act it has been necessary to state the law in Newton's own words, ' Change of motion *is proportional to* the impressed force ', and the force has been measured in gravitational units. The dyne, and, if need be, the poundal, can then be defined, and recognised as not differing in kind from other forces of a more familiar type." 36

The relation $F = ma$ is used. All systems of units are illustrated, but gravitational units predominate in the examples. Experimental work is emphasized.

S. L. Loney, *Elements of Dynamics* (1899).

" The confusion is probably to a great extent caused by the fact that the word ' pound ' is used in two senses which are scientifically different; it is used to denote what we more properly call ' the mass of one pound ' and ' the weight of one pound '. " 37

The relation $P = mf$ is used. All systems of units occur in the examples.

Professor Arthur Morley, *Mechanics for Engineers* (1908).

" . . . . the gravitational system of units has been adopted in the English measures. A serious injustice is often done to this system in books on Mechanics by wrongly defining the pound unit of force as a variable quantity, thereby reducing the system to an irrational one. With proper premises the gravitational system is just as rational as that in which the ' poundal ' is adopted as the unit of force, whilst it may be pointed out that the use of the latter system is practically confined to certain text-books and examination papers, and does not enter into any engineering work. Teachers of Engineering often find that students who are learning Mechanics by use of the ' poundal ' system fail to apply the principles to engineering problems stated in the only units which are used in such cases—the gravitational units. The use of the dual system is certainly confusing to the student, and in addition necessitates much time being spent on the re-explanation of principles, which might otherwise be devoted to more technical work." 38

The relation $F = mf$ is used. Gravitational units are used throughout except Example 10, p. 47. A collection of examples at the end of the book taken from examination questions set for Inter. B.Sc. London (Engineering), A.M.I.C.E., and Board of Education Applied Mechanics does not contain a question relating to C.G.S. or F.P.S. units.

The second part of Quotation 38 requires examination. It illustrates the undesirability of teaching one set of units to the exclusion of all others. Professor Morley complains of the poundal system because he is preoccupied with the gravitational system. His students, who may have been equally preoccupied with the poundal system, may complain of the gravitational system. The beginner should be able to use all systems of units with equal facility. A student who can work in only one system of units is less than half educated. Again, the mechanical or civil engineer cannot expect all early training to be conducted exclusively according to his taste if this runs counter to the requirements of others.

J. W. Landon, *Elementary Dynamics* (1920). (A text-book for engineers.)

" Quite commonly in elementary text-books, the second law of motion is at first summed up in the form force = mass × acceleration, and the student is then given a number of examples to work out, most of which consist in substituting numbers in a formula. This very successfully disguises the true meaning of momentum, and the extraordinary generality of the second law of motion." . 39

" At first sight the remedy would appear to lie in teaching dynamics experimentally, but the author's experience is that this is not so for the majority of the students. The phenomena of everyday life provide innumerable qualitative experiments, and to most students quantitative laboratory experiments in dynamics are neither interesting nor convincing." 40

" In working examples the absolute unit of force has generally been adopted, and, where applicable, the answers have been reduced to units of weight. It matters little, in the author's opinion, whether absolute or gravitational units are used, so long as mass is not defined as weight divided by the acceleration due to gravity. To say that the engineer's unit of mass is 32·2 lbs. is almost to suggest that he is rather lacking in intelligence, and cannot be expected to understand the difference between equality and proportionality. If weight is introduced in the early conception of mass, the student's conception of mass is extremely vague, and his conception of momentum as a physical quantity is even more vague or erroneous. A student who cannot understand the difference in the two units of force, and who has merely to rely on formulae expressed in one particular set of units, is not likely to get any knowledge of dynamics which will be of real use to him." 41

" It is very important to distinguish between MASS and WEIGHT." . . 42

" As we have previously stated, mass is merely the quantity of matter, and is a *scalar* quantity; *weight*, on the other hand, is a *force*, and is a *vector* quantity (its direction being always towards the centre of mass of the earth)." . . . 43

Force is defined as rate of change of momentum. All systems of units are discussed, but of the 100 miscellaneous examples at the end of the book, there is not one in C.G.S. units.

Quotation 39 states that the formula $P = ma$ is often applied mechanically and disguises the true meaning of momentum. Mr. Landon prefers the relation between force and change of momentum to $P = ma$. Will it not happen that if the relation $P = ma$ is ignored, the definition of mass will be successfully disguised and momentum defined in terms of this disguised quantity? Let a student be quite clear on every aspect of the relation $P = ma$ terms, units, experiments, simple illustrations. Let him see pictorial and concrete examples of the application of this important and fundamental relation. Let him mentally work numerous simple direct calculations involving this formula in all systems of units until the use of the relation is automatic and purely mechanical. Then more difficult problems involving thought. The student does not have to think how to dress each morning or how to walk upstairs at night. These quite complicated operations would become embarrassing if thought about—they must be performed automatically and subconsciously. The formula $P = ma$ must be applied with equal facility and readiness.

After this, force may be studied as time rate of change of momentum or space rate of change of kinetic energy. Some problems immediately involve one relation and other problems involve others.

The key to a complete understanding of all the force relations is a thorough comprehension of the idea of mass as embodied in the formula $P=ma$.

There is no reason why the best features of all these books should not be incorporated in all future text-books of elementary dynamics. It is most desirable that all beginners should be trained in the same way. The treatment should be rigorous and philosophical, and can be made interesting by frequent allusion to the history of the subject and the experiments of the founders of its principles—for example, Galileo and Newton (see Quotations 35 and 36). The book should teem with examples of the most varied type—easy, direct, numerical work involving round numbers worked out mentally, more difficult numerical work, graphical work, general problems, data of dynamical experiments, practice in the use of various systems of units. The same treatment, symbols, formulae and equations should be used for all units. Chaotic, mixed, unwanted, unnecessary and irrational units such as the gram weight or gram weight per square centimetre should receive critical examination but should not be used.

REMARKS ON MASS AND WEIGHT BY ENGINEERS

Machinery and Millwork, by W. J. M. Rankine (1876).

" If by the unit of force is understood the weight of a certain standard, such as the avoirdupois pound, then the mass of that standard is $1 \div g$, and the unit of mass is g times the mass of the standard; and this is the most convenient system for calculations connected with mechanical engineering, and is, therefore, followed in the present work." 44

" But if we take for the unit of mass the mass of the standard itself, then the unit of force is the absolute unit; for g is the velocity which a body's own weight, acting unbalanced, impresses on it in a second. This is the system employed in many scientific writings, and in particular in Thomson and Tait's *Natural Philosophy*. It has great advantages in a scientific point of view; but its use in calculations for practical purposes would be inconvenient, because of the prevailing custom of expressing forces in terms of the standard of weight." 45

Applied Mechanics, by John Perry (1899).

" When a body is in motion, it possesses kinetic energy equal to half its mass (its weight in London in pounds divided by 32·2 is its inertia, which is usually, but we think unwisely, called its mass) multiplied by the square of its velocity in feet per second." 46

Calculus for Engineers, by John Perry (1897).

" To satisfy the College men who teach engineers I would say that ' The unit of Mass is that mass on which the force of 1 lb. produces an acceleration of 1 ft. per sec. per sec." 47

"We have no name for unit of mass, the engineer never has to speak of the inertia of a body by itself." 48

Steam Power, by W. E. Dalby.

In the chapter on the balancing of reciprocating masses the formulae $F=MA$ poundals and $F=MA/g$ lbs. wt., $M=$mass in pounds are given. 49

The Steam and other Heat Engines, by J. A. Ewing.

"Let M be the mass of the piston, piston rod, and cross-head in pounds, and a its acceleration at any instant in feet per second per second, the force required to accelerate it is Ma/g, in pounds weight . . . ." 50

Theory of Machines, by Louis Toft and A. T. J. Kersey.

The mass of a body is a measure of the quantity of matter in it, and its value in engineer's units is the ratio of the weight of the body in pounds, at any part of the earth's surface, to the value in feet per second per second of g, the acceleration of a body falling freely under gravity at the same place. The weight W lb of a body is the pull of the earth on the body, and varies slightly with the locality, but the ratio W/g is constant.

The unit of mass is then the quantity of matter in a body weighing g lbs." 51

Later the formula $F=Ma$ is given.

The Quotations 44 to 51 occur in well-known text-books of engineering. Not one of these writers avoids the use of the term mass, but all use the pound weight as the unit of force. The use of the pound as the unit of mass as well as the unit of force is the cause of all the confusion and difficulty of beginners. Quotations 47 and 48 avoid this inconsistency of units, but reveal the astounding fact that engineers habitually use a unit without giving it a name. It must also be remembered that although mechanical engineers may not speak frequently "of the inertia of a body by itself" there are aeronautical, civil, electrical, marine and other types of engineer, quite apart from those who are not engineers. All students receive their early training together, and the subjects of physics and mechanics must be taught not in terms of engineering but in the best interests of science and engineering.

If m is mass in pounds, $F=ma/g$ is true numerically but not dimensionally.

If W is weight in pounds, $F=Wa/g$ is true numerically and dimensionally. It is true of a body at a place where it has little or no weight: it is true of bodies everywhere. It is true of objects having no weight. It is the formula $F=ma$, but m is expressed as the ratio of two variables. Why is the ratio of two variable quantities W and g used in preference to a single constant quantity m?

More than one reason for this may be given:—

(1) W is introduced for popular convenience.

Surely, precise scientific expression should not be sacrificed for popular convenience.

(2) W is introduced because engineers think of force in pounds weight.

All the more reason for expressing mass in terms of a unit consistent with force in pounds weight. With force in pounds weight, mass should be in slugs.

(3) W is used to represent pounds mass.

This procedure is really bad and dimensionally incorrect. The use of chaotic units in a chaotic formula should be rigorously rejected. If a problem involving pounds mass as well as pounds weight arises, one of these units must be changed before substitution is made in any rational formula.

REMARKS ON MASS AND WEIGHT BY MATHEMATICIANS

Horace Lamb, *Dynamics* (1914).

In Chapter II there are sections headed *Gravitational Units* and the *Absolute System of Dynamics*. Both equations $P=ma$ and $P=Wa/g$ are given.

" It is now usual to designate inertia, when regarded as a measurable quality, by the term ' mass '; the mass of the body on the present reckoning is therefore W/g." 52

" It is scarcely necessary to insist on the special sense which the word ' force ' has come to bear in Mechanics. It is perhaps unfortunate that some more technical term was not introduced instead of a word which, in popular language, has so many different meanings. The usage is, however, long established and must be accepted." 53

" There is not the same agreement, although there has been much controversy, as to the use of the word ' weight '. In ordinary language this is employed in a great variety of senses. Thus it may mean the actual statical *pressure* which a body exerts on whatever is supporting it, as when we speak of the ' weight ' of a burden; it may mean the *ratio* which this pressure bears to that exerted by a pound or a ton; it is often used virtually in the sense of *mass*, as when we refer to the weight of a projectile; when, again, we speak of the ' weight ' of a blow the idea is (vaguely) that of *momentum*. The one sense in which the word is *never* used in popular language is that of the gravitational attraction on a body. This is, of course, equal to the statical pressure above referred to, but it is not identical with it; it is a force exerted *on* a body, not *by* it. Unfortunately, this new and alien sense is precisely that which some writers of eminence have sought to attach exclusively to the word ' weight ' in Mechanics. In the author's opinion it is best not to attempt to specialize altogether the meaning of so familiar a word, but to use it freely in whatever sense may be convenient, whenever there is no risk of misunderstanding. When there is danger of confusion, some other term, such as ' mass ' or ' gravity ', may be employed to indicate precisely the sense which it is wished to convey." 54

" In the rest of this treatise we follow the absolute system, as by far the most convenient for general application." 55

A. E. H. Love, *Theoretical Mechanics* (1906).

" It is clear that the definition of mass by means of mutual action is more general and more fundamental than that by means of weighing. We shall show . . . . that the determination of masses by weighing is a particular case of the determination by means of mutual action.

" Since we are accustomed to estimate the *quantity of matter* in a body by weighing the body, it is customary to state that the quantity of matter in a body is equal to the mass of the body." 56

" Thus the mass of the body provides a measure of its *inertia*." . . . 57

Alfred George Greenhill, *Hydrostatics* (1894).

" With the Gravitational Unit of Force, the weight of a body is at once the measure of the quantity of matter in the body, and also the force with which it is apparently attracted by the Earth; and the word Weight may be used in either sense without ambiguity or confusion, when dealing with hydrostatic problems on the surface of the Earth." 58

Alfred George Greenhill, *Notes on Dynamics* (1908).

" We use the word Weight advisedly, so as to agree with the precise language of the successive Acts of Parliament.

" The weight (poids, pondus) of a body is the quantity determined by weighing; to weigh (poise) the body, it is placed on the one of the scales of a balance, and equilibrated with certain lumps of metal in the other scale, called weights; and the sum of these weights is called the weight of the body." . . 59

" A force of F pounds acting on a weight of W lb. will give it acceleration a f/s^2, such that $a/g = F/W$." 60

Professor E. T. Whittaker, *Analytical Dynamics* (1937).

" If any set of mutually connected particles are in motion, the acceleration with which any one particle moves is the resultant of the acceleration with which it would move if perfectly free, and acceleration directed along the lines joining it to the other particles which constrain its motion. Moreover, to the several particles A, B, C, ..., numbers m_A, m_B, m_C, ... can be assigned, such that the acceleration along AB due to the influence of B on A is to the acceleration along BA due to the influence of A on B in the ratio $m_B : m_A$. The ratios of these numbers m_A, m_B, ... are invariable physical constants of the particles.

" The evidence for the truth of this statement is to be found in the universal agreement of the calculations based on it, ... with the results of observation.

" It will be noticed that only the *ratios* of the numbers m_A, m_B, m_C, ... are determined by the law; it is convenient to take some definite particle A as a standard, calling it the ***unit of mass***, and then to call the numbers m_B/m_A, m_C/m_A, ... the ***masses*** of the other particles m_B, m_C, ...

" The mass of the compound particle formed by uniting two or more particles is found to be equal to the sum of the masses of the separate particles. Owing to this additive property of mass, we can speak of the mass of a finite body of any size or shape; and it will be convenient to take as our *unit of mass* the mass of the 1/1000th part of a piece of platinum known as the *standard kilogramme*; this unit will be called a *gramme*, and the number representing the ratio of the mass of any other body to this unit mass is called the *mass of the body in grammes*." 61

" In general, if an acceleration represented by a vector f is induced in a particle of mass m by any agency, the vector mf is called the *force* due to this cause acting on the particle; and the resultant of all the forces due to various agencies is called the *total force acting on the particle*." 62

Professor Karl Pearson, *The Grammar of Science* (Part 1, 1911).

" The proper measure of mass is found to be a ratio of mutual accelerations, and force is seen to be a certain convenient measure of motion, and not its cause. The customary definitions of mass and force, as well as the Newtonian statement of the laws of motion, are shown to abound in metaphysical obscurities." . . 63

Quotation 54 refers to the many popular meanings of the word " weight ". In a beginner's science course, the earliest opportunity should be taken to explain the scientific meaning of this term. The beginner is quite familiar with the *idea* of mass but not with the *word*. It is not difficult to put this right, but the difficulty of correction becomes greater and greater if the young pupil is allowed to establish the habit of using incorrect terms in a scientific problem. This also gives the opportunity of discussing the more difficult idea of " weight ", the gravitational attraction of the earth. This word must not be used in any other sense. The mind of the young beginner is easily moulded. Bad habits must be eradicated at once and the process of correction must be immediate. Any delay increases the difficulty of correction and the possibilities of confusion. The distinction between mass and weight must be discussed at the very beginning. Once these terms have been explained, they must receive careful and persistent attention.

Quotation 58 studiously avoids all reference to mass. The word does not occur anywhere in *Notes on Dynamics*. The word " weight " is used for " mass " throughout the book. The book contains no force or work problems in C.G.S. or F.P.S. units. It is probable that Prof. Greenhill originated the use of the formula $F/W=a/g$. This formula involves a different unit of weight in each place.

Quotation 61 is a statement that could scarcely be improved upon. It is simple, lucid, and philosophically rigorous. It is as easy to realize that $m_1/m_2=f_1/f_2$ as $P/W=a/g$, but the first relation is a ratio of concrete quantities and is true universally, whereas the second relation involves the ratio of two forces in different directions, and these forces must act on the same body at the same place relative to the earth.

Quotation 63 indicates the futility of philosophical discussion, at any rate for beginners. Quoting from J. B. Stallo's *Concepts of Modern Physics*:—

" Mass and motion being mutually inconvertible, mass is absolutely inert." 64

" Once more, then, science is in irreconcilable conflict with one of the fundamental postulates of the mechanical theory. Action at a distance, the impossibility of which the theory is constrained to assert, proves to be an ultimate fact inexplicable on the principles of impact and pressure of bodies in immediate contact. And this fact is the foundation of the most magnificent theoretical structure which science has ever erected—a foundation deepening with every new reach of our telescopic vision, and broadening with every further stretch of mathematical analysis." 65

Action at a distance, variable mass, absolute space and time, the real object, and other similar problems are not suitable for elementary discussion.

REMARKS ON MASS AND WEIGHT BY PHYSICISTS

W. Thomson and P. G. Tait, *Natural Philosophy* (1879).

" The Quantity of Matter in a body, or, as we now call it, the Mass of a body, is proportional, according to Newton, to the Volume and the Density conjointly. In reality, the definition gives us the meaning of density rather than of mass; . . . ." 66

" Newton further states that a practical measure of the mass of a body is its Weight. His experiments on pendulums . . . ." 67

J. Clerk Maxwell, *Matter and Motion* (1882).

The Chapter on Force is particularly illuminating. See also Chapter IV, Maxwell's *Heat*.

" . . . . Of these three methods, that of weighing depends on the attraction between the acid and the earth, that of measuring depends on the volume which the acid occupies, and that of titration depends on its power of combining with potash." 68

" In abstract dynamics, however, matter is considered under no other aspect than as that which can have its motion changed by the application of force. Hence any two bodies are of equal mass if equal forces applied to these bodies produce, in equal times, equal changes of velocity. This is the only definition of equal masses which can be admitted in dynamics, and it is applicable to all material bodies, whatever they may be made of." 69

J. H. Poynting and J. J. Thomson, *Properties of Matter* (1905).

" This constancy of acceleration under a given force is expressed by saying that the mass of the body is constant." 70

" Further experiment shows that the acceleration of a given body is proportional to the force acting on it." 71

"The masses of bodies are proportional to the forces producing equal accelerations in them." 72

"But for scientific purposes all over the world the unit of mass is the gramme, the one-thousandth part of the mass of the piece of platinum-iridium called the Kilogramme-International, which is kept at Paris." 73

Henri Poincaré, *Science and Method*, translated by Francis Maitland. No date (after 1913).

Poincaré has made contributions to philosophy, mathematics, astronomy and physics. The section in the book on *Mechanics and Radium* is of interest. The following quotation is from his remarks on the *New Mechanics and Astronomy*. He discusses real and fictitious mass, but the only bearing on the present discussion is the fact that a body may have no weight, but it may have mass, and mass varies with velocity, as in the case of a rapidly moving charge.

"Mass may be defined in two ways—firstly, as the quotient of the force by the acceleration, the true definition of mass, which is the measure of the body's inertia; and secondly, as the attraction exercised by the body upon a foreign body, by virtue of Newton's law. We have, therefore, to distinguish between mass, the coefficient of inertia, and mass, the coefficient of attraction." . . 74

See also Routh's *Analytical Statics*, Vol. II.

Quotations 66 to 74 indicate remarkable uniformity in point of view and treatment. The physicist uses $F=ma$, a formula which may be applied without modification to any system of units.

John Cox, *Mechanics* (1904).

The treatment in this book approximates closely to that advocated in this paper. It is not altogether an elementary text-book but one for university classes. All systems of units are discussed and included in the examples. Beginners require more numerous examples of great diversity of type and difficulty.

" . . . . $P=ma$.

"The simplest way of solving dynamical problems is to use this formula in conjunction with the Kinematical formulae . . . ." 75

The following quotations are from the Preface:—

"It is a common complaint that though the principles of Mechanics are the simplest and the earliest to be discovered in the whole range of Science, and, moreover, are directly illustrated in almost every act of our lives, more difficulty is found in giving beginners a real grip of them than with any other branch of Physics.

"This I attribute largely to the way in which the text-books deal with the subject. The student usually opens the book upon a chapter in which such leading concepts as matter, force, mass, particle, rigid body, smooth body are treated in definitions of a line or two each, before he sees any reason for their introduction at all. He is probably warned that philosophers are not agreed about the nature of matter; that motion is purely relative; that force is a misleading idea borrowed from our muscular sensations and better got rid of; and that no such things as mathematical particles, rigid bodies and smooth bodies exist in nature. He naturally concludes that Mechanics is an abstruse subject having nothing to do with realities or common sense." 76

"The second chapter plunges him into the mathematical study of motion in the abstract. . . . . To his previous confusion he adds the conviction that this is only another branch of the pure mathematics he has hitherto found so little use for." 77

"At last there is a chapter on the Laws of Motion, so inadequately treated that he often ends by believing that they were made up by Sir Isaac Newton, the author, so far as he is aware, of the whole subject. The rest of the book is too often merely geometrical and trigonometrical gymnastics." 78

"In recent years many text-book writers have attempted to break away from this mischievous tradition. Some have tried to rewrite the whole subject from the latest point of view of Energetics. But this is surely to begin at the wrong end. According to the biologists the bodily development of the individual is an epitome of the development of the race. Is not this a hint that the historical method is the natural way of attacking a subject of study? Others have sought to discard the idea of force, and speak only of mass-accelerations. *Naturam expellas furca.* It is rarely indeed that they manage twenty pages without getting back to the old point of view. With proper caution the use of this concept is as valuable as it is historically right and inevitable. Still others have set the student to rediscover the subject for himself by experiment. But this wastes too much time on mere manipulation, and leaves the student's knowledge in mid-air, unrelated to all that has gone before him in the course of actual discovery. It seems a pity that he should close the book without a glimmering of personal interest in his predecessors, the great investigators, and forego the insight into philosophic and scientific method which a study of the development of mechanics evokes insensibly and unawares." 79

"After learning and teaching Mechanics for ten years on the traditional system described above, I was called on, as a lecturer under the Cambridge University Extension Scheme, to explain the principles to audiences without any previous mathematical training, but often composed of engineers, plumbers, and other workmen who had derived excellent practical notions on the subject from their experience. Obliged thus to recast the subject in my own mind, I found it possible to present all the main principles with the aid of ordinary arithmetic and the simplest geometrical diagrams. At this stage Sir Robert Ball's admirable lectures on *Experimental Mechanics* gave me great assistance. My experience with these popular audiences reacted with advantage on my teaching with classes in the university, and fired me with the ambition to write a text-book on Mechanics. But a sight of Sir Oliver Lodge's excellent *Mechanics* in Chambers' series put an end to this wish for a time. Some ten years ago I stumbled on the first German edition of Professor Mach's *Die Mechanik in ihre Entwickelung*. I am ashamed to say that this fascinating book was my first introduction to the historical development of a subject I had taught so long. Since then my teaching has been based more and more on the lines laid down by Mach, and as I have found it impossible to induce ordinary students to read the original, even when translated, I recurred to the idea of writing a text-book which should yet be based on Mach's method." 80

" Until Mechanics is clad in its historical flesh and blood, it will remain the dull and tiresome subject that has convinced so many generations of students that an abysmal gulf separates theory from practice." 81

The following works may be consulted for further historical detail:—

Dr. Charles Singer, *Studies in the History and Method of Science*.

Prof. Wolf, *A History of Science, Technology, and Philosophy*.

APPENDIX 2

This paper was written before the author was aware of a *Report on the Teaching of Mechanics in Schools*, prepared for the Mathematical Association (1930, published by G. Bell & Sons, Ltd.).

The Sub-Committee which prepared the Report was composed as follows:—

W. J. Dobbs.
W. C. Fletcher.
C. J. A. Trimble.
Miss D. R. Smith.
Miss L. M. Swain.

A. Robson, *Secretary*.
C. O. Tuckey, *Chairman*.

The Report is a valuable contribution to the teaching of Mechanics. The parts of it relating to the present discussion occur chiefly in paragraphs 1, 2, 3 and 6. Unfortunately, the assistance which science and mathematical masters might give to each other is ignored. No attempt is made to explore possibilities of co-operation (see Appendix 4), and the report deals primarily with the teaching of mechanics as a part of elementary mathematics for the School Certificate Examination or for non-specialists.

" Mechanics . . . . is sometimes included in science and sometimes in mathematics. If it is included in science as part of physics, it is generally done very slightly, the main attention being given to heat, light and electricity. The reason for this is probably that the science master finds mechanics less suitable than the other subjects for his characteristic method of teaching, with its constant appeal to experiment." 82

"In mechanics . . . it is to experience rather than to experiment that constant appeal must be made, though it is not denied that experiment may be of great value. Moreover, the very large amount of written example work desirable is more consonant with the atmosphere of the mathematical class-room. Hence it is more usual to take mechanics as part of mathematics." 83

"Probably the ideal arrangement would be that the same man should teach the same group of boys in both mathematics and physics; he would then be able to use experiment without the difficulties felt by the mathematical master, and at the same time to do the bulk of the work in mechanics in the mathematical hours.

"This arrangement, however, is not often found to be possible, and, as things stand, this Committee has no hesitation in endorsing the correctness of the usual procedure by which mechanics is included as part of mathematics." . 84

These views are narrow and short-sighted. Mechanics is taught both by science and mathematical masters. Who teaches it, and whether it forms part of mathematics, or physics, or both, is not relevant unless the principles of mechanics are different in the different subjects. It is not for the science master or his mathematical colleague to prescribe for the other how much experiment, experience, geometry, trigonometry, etc., he should bring into his teaching. Nor should there be controversy as to right of way. The fullest co-operation between masters and the closest correlation of subject matter and methods are essential in the best interests of the student. What is really unfortunate is that the spirit of Quotation 84 dominates the recommendations of the report.

I am strongly of the opinion that statics and dynamics should form part of mathematics and that statics, dynamics, and hydrostatics should form part of physics. The same treatment, notation and formulae of fundamental principles should be followed by both mathematical and science masters.

"The bulk, therefore, of this *Report* is devoted to a discussion of the best methods of carrying through a comparatively limited pre-certificate course of mechanics, though the needs of specialists and other more advanced students have also been considered." 85

The needs of the non-specialist are fully provided for in a course of General Science. These boys will not as a rule proceed to the stage at which trigonometry is part of the mathematical course.

"Mathematicians, scientists, engineers, and others need not be separated for teaching purposes on any other basis than that of ability, unless it happens that, e. g., specialists in physics and chemistry have insufficient time to work with the mathematical specialists." 86

It will be agreed that all boys who study mechanics as part of mathematics will receive the same training together throughout the school.

Paragraphs 2 *and* 3

These paragraphs deal with experience and experiment.

" The purpose of observation or experiment should be to enable boys to realise how the thing works. Quantitative results are relatively unimportant." 87

" Perhaps the whole may be summed up by saying that, while quantitative work may be dispensed with, effective steps can and must be taken to meet the weakness of geometrical imagination, which is one of the great stumbling blocks." 88

Many will neither entirely agree nor entirely disagree with Quotations 87 and 88. I should regard Galileo's experiments on uniform acceleration as important although quantitative, and if I taught mechanics as part of mathematics there would be many uses for the experimentally determined values of s and t. I regard the evaluation of g by all manner of different methods as an extravagant use of valuable time.

Paragraph 6

" This section of the report deals with the most controversial parts of the subject: how and when to introduce the idea of mass; whether or not to use an absolute system of units; if such a system is used, whether it should be based on the unit of mass or the unit of force.

" These and allied questions are the questions most hotly debated among teachers of mechanics. The sub-committee preparing this report has given a great deal of time to their consideration, and if complete unanimity has not been reached, at least the opposing views were found at the end to differ much less widely than at the beginning of the discussions." 89

The report contains an excellent account of the controversial points. There was complete agreement of the Committee on two matters:—

6.1. " *Force and rate of change of momentum should be thought of by the learner as different things.*" 90

6.2. " *The double meaning of the word ' pound ' should be introduced early.*" 91

" Much of the difficulty, real or supposed, of dealing with mass would be abolished if teachers would adopt a very simple precaution.

" The word ' pound ' is used in mechanics in at least two senses, if not by pure mathematicians certainly by engineers. This may be regrettable, but we cannot change it. The pedagogic error consists in starting with the one sense (force) without warning that there is another (mass), and that it is the other which is fundamental. The further the subject is pursued with the one meaning of the term, the more deeply rooted does that meaning become, and the greater the difficulty in making a change. In statics we measure forces in pounds and sow the seed of trouble. Frequently we continue the practice on starting dynamics, and the association of pound with force gains strength. Still further, a new term which is coming into use, the ' second-pound ' as a unit of impulse or momentum, also tends to fix the notion of pound as force if a careless treatment is adopted. The same objection does not apply to sec. lb. wt.

" After all this we attempt a ***volte face*** and teach that a pound is not properly a force at all, but something quite different—a mass." 92

I should like to have quoted the whole of this paragraph. It very ably exposes the inevitable confusion arising from not distinguishing between mass and weight.

It will be observed that the student of mechanics has a further treat in store for him—the second-pound as a unit not only of impulse but also of momentum.

" While we do not need ' mass ' in statics and while long established custom sanctions the use of ' pound ' as a unit of force, it is strongly urged that from the outset the other meaning of ' pound ' should be kept in mind so that, when occasion arises for the use of that meaning, boys' minds should not be unprepared and resistant." 93

Surely this can only mean that mass should be considered from the outset. It is stated in the report that complete agreement was reached by the Committee with regard to 6·2. A majority of the Committee were in agreement with 6.3, 6.4 and 6.5, but 6.3 is inconsistent with 6.2.

6.3. " *In teaching the average beginner the fundamental equation of dynamics should first be obtained and used in the form $P/W=f/g$, the consideration of mass being postponed for a time.*" 94

" Dynamics deals with the production of motion in matter by force. The property of matter which concerns us here is mass." 95

Quotations 94 and 95 seem to suggest that, in the case of a beginner, the fundamental dynamical relation (which deals with mass) should be dealt with by postponing the consideration of mass.

" As set out in detail later, it is less difficult than is sometimes supposed to explain how mass is measured, but the difficulty is for the beginner to acquire any feeling of the significance of the mass as distinct from the weight of the same lump of matter; the same standard pound provides the unit in each case, either by its mass or by its weight, and the two numerical measures are therefore the same." 96

" Fortunately it is possible to make the practical question involved—the question of the motion produced by a given force—essentially simpler for the beginner by the postponement of the consideration of mass, and the use of gravitational instead of absolute units." 97

" In this method the fundamental equation of the subject is not used in the form $P=mf$ . . . . Instead it is obtained in the form $P/W=f/g$.

" This last fact affords the additional advantage that the forces P, W, and also the accelerations f, g, may be measured in any units we please." 98

In connection with Quotation 94 it may be pointed out that there is no difficulty with regard to a beginner using the idea of mass, particularly if the distinction between mass and weight has been dealt with at an early stage in the science course. If this has been done the use of $P/W=a/g$ is likely to lead the student into utter confusion. The application of this formula to gravitational units cannot fail to give correct results, particularly and only in chaotic units. The user of this formula not only condones the chaotic units pound weight and pound mass, but he proceeds to inject chaos into rational systems of units. This formula is a complete failure in connection with C.G.S., F.P.S. and the true engineers' system of units, in which the unit of mass is the slug. It will be observed that in order to introduce this formula, every idea of rational scientific procedure is sacrificed. Still further, the whole scheme must be scrapped in the end, so that at some later stage the student must change over to rational methods. The user of this formula proceeds to use it for his own purposes and, having created a condition of chaos so far as others are concerned, these unfortunate others and the pupil himself are left to extricate order out of confusion.

To begin with, this formula involves:—

1. Postponement of the idea of mass until a rational system of units comes into use.
2. The use of chaotic units exclusively.
3. The use of chaotic formulae.
4. The introduction of the idea of weight into problems in which weight is not directly involved.
5. A change-over to a rational system after training the student in chaotic units and formulae.
6. Complications with regard to quantities such as moment of inertia (see Quotations 14 and 15).

The remedy for all this confusion is very simple:—

1. Deal with mass early.
2. Call the engineers' unit of mass a " slug ".
3. Never use chaotic units or formulae. If chaotic units arise in a problem, change them to rational units.
4. Work problems in all systems of units.
5. Use $P=ma$, from which unit force produces unit acceleration in unit mass.
6. Co-operative effort by mathematical and science masters, so that there is complete uniformity of treatment and it is not necessary to change any of the earlier ideas.

6.4. *Mass.*

The report indicates an inferiority complex with regard to the successful explanation of mass to beginners. Teachers of mechanics should find that this has already been carried out by the science master unless the latter finds that all his early hard work is regularly shattered by the use of $P/W=a/g$.

"The difficulty that arises from the use of the pound and gram both as units of mass and units of force is inevitable, but is not so great as it is sometimes thought to be." 99

Why create difficulties by introducing unwanted units? If a student is trained to use rational units rationally there are no difficulties, even in the training of the student.

6.5. *The change from the habitual use of gravitational units to that of absolute units.*

"Those who commence with the use of gravitational units should, if they continue their studies sufficiently far, change over to the habitual use of absolute units (except for working certain types of numerical examples)." . . . 100

"It is recommended that the change should be effected in two stages: a preliminary stage in which absolute units are introduced and used for the first time while gravitational units still remain dominant, and a final stage in which absolute units become habitual and gravitational units are merely reserved for problems to which they are specially suited." 101

Comment seems superfluous. The procedure here advocated is a fitting climax to the use of $P/W=a/g$.

6.6. *The "Absolute" Theory.*

A minority of the Committee were in agreement with 6.6.

"The last three sections represent the views of what may be called the gravitational school; this section attempts to state once more the views of the other school, dominant at the end of the last century, but now gone out of fashion so far as school-books are concerned." 102

"The essential points in the views here set forth are the introduction from the outset of the term mass in its simple sense and the correct use of the word pound." 103

". . . . shall we use absolute units from the start or shall we use gravitational units at first and only later, and gradually, introduce the others?" . . . 104

"It may be noted that mass is a more primitive thing than force in the sense that it is much easier to define a standard of mass than a standard of force; . . . ." 105

". . . . dynamics is essentially the study of the relations of three entities . . . force, matter, and motion, and any treatment which ignores or obscures one of the three is essentially defective." 106

May I add:—

1. Use all units, but all formulae and working should require the conversion of chaotic or irrational units to a rational system before introduction into dynamical relations.

2. Use $P = ma$ to define the units of force or mass in any rational system of units.

3. The early distinction between mass and weight should be emphasized by the science master, and the mathematical master should continue a treatment already begun. The science master will also continue at a later stage, and there should be no modification in the method of working at any stage of the student's career.

4. Reject such units as gram weight or grams weight per square centimetre.

5. Call the engineers' unit of mass a slug.

APPENDIX 3

" WEIGHT " MEANING " MASS "

There must be something radically wrong with any system of education which leads to a necessary distinction and difference between science and popular science. Similarly, the popular incorrect use of the word *weight* for *mass* is an indication of defective education. The scientific worker has been far too tolerant of the superior attitude of those who criticize a split infinitive but connive at and condone the inaccurate use of common words. In my opinion it is worth while making an attempt to correct popular speech and thought. This could be done in a single generation. The wording of Acts of Parliament should be corrected, and not tolerated. Teachers of English and mathematics should begin early to correct this slovenliness and inaccuracy of speech. Science masters do not as a rule teach *all* students, whereas teachers of English and mathematics not only influence the whole rising generation, but they can make that influence felt early and before the science master gets a chance.

I have expressed the opinion that the misuse of certain words could be corrected in a single generation. Lord Woolton could go far in this direction by suitable wording of food edicts. If the cigarettes called "Weights" had been called "Masses" the public might become curious, but if the makers of these cigarettes would change the name and state the reason for the change on the packet or in newspaper publicity, this would be a public service. Similarly, if sweets, sugar, meat, etc., were sold by mass instead of weight, the word mass would rapidly take the place of weight, and with its correct meaning.

What is implied in Paragraph 6.4, Appendix 2, " the teaching of mass being postponed for a time " ?

There are two unsatisfactory suggestions here:—

1. A student is to be allowed to use the word " weight " for " mass " until a later stage of the work is reached when mass calls for consideration. Does this not encourage inaccuracy and slovenliness of thought and speech ?

2. A student beginning dynamics is to avoid the consideration of mass. This implies that early studies in dynamics should avoid the consideration of dynamics and seek some alternative method of study in which fundamentals are

side-tracked. This is not dynamics but subterfuge. It leads only to unfailing accuracy in the substitution of numbers in a formula. It would be far better to defer the teaching of dynamics until a student is capable of studying the subject in a rigorous and proper manner (see Quotations 38 and 41).

APPENDIX 4

The Report of Sir J. J. Thomson's Committee on Natural Science in Education (1918) contains the following:—

" Much stress has been laid in the evidence before us on the need for correlation of school teaching in Mathematics and Natural Science. There is evidently a strong feeling, both among school teachers and among those who have to deal with the products of the secondary schools at the universities and in the industries, that both subjects would gain if they were taught with more reference to each other, and, so far as possible, by the same teachers. At present it is easier to get a science master who is able to teach elementary mathematics than a mathematical master with a corresponding knowledge of science; but few schools can spare a science master for any part of the mathematical work. The need for co-operation is not felt only on the side of the natural sciences; the Committee of the Mathematical Association express the views that effective correlation between the two subjects is rare and is not increasing, that an ideal arrangement would assign the teaching of mathematics and of physics largely to the same masters, and that the education of teachers of mathematics should be conducted with this end in view." 107

" Any want of co-ordination between Science and Mathematics has a particularly unfortunate effect on the teaching of mechanics. This subject is rightly regarded as of the greatest importance. It is the basis of most parts of physics, and most schools make some provision for teaching it, either as a branch of Mathematics or as an experimental subject, or from both points of view; but the evidence we have received indicates that in a considerable number of schools the results obtained are far from satisfactory. This is much to be regretted, for the subject is especially suitable for training the student to use his mathematics and to apply it to the problems which he may have to face. The power of doing this is a very valuable asset for those who will become engineers or be occupied with other applications of Science to industry; while, apart from this aspect, it affords a mental training of the highest importance and gives to many an interest in mathematics and a grip of its principles which, without it, they would never acquire." 108

I am of the opinion that a regrouping of subjects is desirable in framing the school curriculum. This could be divided into four main divisions—English, Languages (ancient and modern), Science, Aesthetics. Science could be sub-divided into Mathematical Science (Mathematics, Physics, and Astronomy and Natural Science (Biology, Chemistry, and Geology). In this case, mathematics and physics would be one subject and there could be considerable pruning of both elementary mathematics and elementary physics. The syllabuses of elementary examinations require far too much detail. A higher standard of fundamentals would be a better foundation for the future.

EPILOGUE

The pound weight is the unit of force used by engineers. Why do some teachers of mechanics develop the subject as though the pound mass is the engineers' unit of mass? This not only embarrasses engineers, but it gives rise to utter confusion in the use of other units and results in duplication of formulae.

There should be drastic and complete eradication of slovenliness in the use of words, units, formulae and an insistence on strict rigour of treatment. All teachers of dynamics should develop the idea of mass, and students should work problems in all units with equal facility.

All this can be achieved very simply, as follows:—

1. The terms mass and weight should, from the very beginning, be used carefully and correctly, with their accepted scientific meanings. The legal terminology of Acts of Parliament should be ignored.

2. A name is required for the gravitational unit of mass. Professor Worthington suggested the name slug.

3. Use appropriate standard notation and descriptions:

$P=ma$ dynes, poundals, or pounds weight.
Momentum $=mv$ gm.-cm./sec., lb.-ft./sec., slug-ft./sec.
Kinetic energy $=\frac{1}{2}mv^2$ ergs, ft.-pdls., ft.-lbs. weight.
Potential energy due to height $=Wh$ ergs, ft.-pdls., ft.-lbs. weight.
Moment of inertia $=mk^2$ gm.-cm^2, lb.-ft^2, slug-ft^2
Pressure $=P/A$ dynes/cm^2, pdls./ft^2, lbs. wt./ft^2
Density $=m/V$ gm./c.c., lbs./c. ft., slugs/c. ft.

4. Use the same formulae and treatment for all systems of units and rigorously reject chaotic units. No duplication of formulae will arise.

APPENDIX 5

TREATMENT OF THE FORCE EQUATION IN AMERICA *

The practice in America appears to be to teach the conception of mass more generally than in this country and, in consequence, the equation $F=ma$ is more universally used. There appears to be the same confusion of units and the same difference of opinion with regard to whether the poundal should be used or abolished. The slug appears to be more generally known than in this country.

---

* The whole of this article on mass, with the exception of Appendix 5, was written before consulting the American journals.

Another system of units is finding favour in certain quarters—the metre-kilogram-second system or M.K.S. system, in which the unit of force is the newton. 1 newton $= 10^5$ dynes.

Reference may be made to an article in the *American Physics Teacher* (August 1939, **7**, no. 4), "On the Meaning of a Constant in a Physical Law", by Prof. H. M. Dadourian. The article takes the form of a discussion between Simplicio, Sagredo, and Salviate. The topic for discussion is "the question of re-forming the equations of mathematics and the laws of physics along the lines laid down by certain authors of elementary physics text-books, who write the force equation as $F=kma$".

Most of the discussion relates to showing that k in this equation is an unnecessary constant of conversion. The only necessary constant in the equation is m, which is a constant of proportionality, which has dimensions, and which is a physical property of the body whose motion is considered.

The remarks on the poundal and the slug are vague and unsatisfactory and do not fit well into the argument.

> "SIMP. I grant that your objections against conversion constants are well taken, so far as simplicity of equations and consistency of notation are concerned ; but I maintain that there is something to be said for k in $F=kma$, for the poundal and for the slug on practical and pedagogic grounds. These enable students to solve problems that involve weight. I tell you it works.
>
> SALV. There can be no practical use for them ; otherwise they would be used in engineering and advanced physics. Who has ever heard weight spoken of as so many poundals, or so many slug-feet-per-second-per-second, outside of a classroom in elementary physics ? Even in such classrooms these gewgaws are not mentioned after the first chapter or two. As for the pedagogic argument that 'it works', it works simply because problems are so worded that 'it' can't help working".

Later, Salviate makes the following suggestions :—

1. We can refrain from using the poundal and the slug.
2. We can refrain from using a superfluous k whose value jumps back and forth with values 1/981, 1/32, and 1.
3. We can stop talking about the "pound-force" or the "force-pound" and the "pound-mass" or the "mass-pound".

He now prescribes a remedy :—

> "Let me show you a way which is not only theoretically sound but also very simple. Train the student to follow these rules :—1. Using letters, find an expression for the required quantity. 2. Perform the necessary literal substitutions until every letter in the expression represents a given or a known quantity. 3. Put the given numerical values in the expression after the second rule is complied with—never before. These are invaluable rules ; they save time, reduce the number of numerical blunders and are aids to clear thinking. As an illustration of the simplicity of this method, suppose w and a are given, and F is required. The student writes $F=ma$; then, since m is not given, he looks for an equation by means of which he can eliminate it ; he thinks of $w=mg$ and making use of it, obtains $F=wa/g$. The rest is a matter of mere substitution of numerical values. As simple as that."

The rules for solving a problem here prescribed completely beg the question. Students have no difficulty with literal formulae—the mistakes and the confusion in the solution of force problems arise entirely when numbers are put into the formulae. The poundal and the slug having been abolished, what numerical substitutions are necessary?

"SIMP. Then you favor writing w/g for mass, and $F/w=a/g$ or $F=wa/g$ for the force equation.

"SALV. Not at all. In fact, I object to these more than to $F=kma$, for the following reasons. The equation $F/w=a/g$ reduces a physical law, a very important one at that, to a mere numerical equation; for both F/w and a/g are pure numbers. On the other hand, $F=(w/g)a$ suffers from the disease of circularity, when considered as the definition of force. You will admit that a definition which expresses a force F in terms of another force w is not a legitimate definition. As to the use of w/g as the general definition of mass, there is little to be said for it. Authors who represent mass as w/g in the equations of physics ignore the fact that mass, being the most important property of matter, deserves a symbol of its own; they confuse a physical quantity with a particular way in which it may be measured under certain limited circumstances; they revert to geocentricism without being aware of it."

In the *American Journal of Physics* (April 1940, **8**, no. 2) there is a note by Professor W. W. Sleator with which I entirely agree. He says:—

"I should have agreed with Salviate in his defining equations, and in his definition of force in terms of mass and acceleration. But I should have objected to the condemnation of the poundal, and his bringing together 'the poundal and the slug' as if they both arose from the same objectionable equation. . . . He makes out an unfair case against the poundal, and by implication an unfounded charge of perversity or stupidity against those who use it. . . . Yet Dadourian will have none of the slug. Neither will he write $m=w/g$. He then has no English unit of mass at all. . . . . Dadourian's general contention seems to be that we should use an absolute system of units, but, in order to conform to the usage of practical men, we should disguise the absolute unit in the result and express forces in pounds. This is one more variation on the common practice of placating the engineer by not trying to teach him anything. . . . We all teach a system of grams, centimeters, dynes, and ergs. To these correspond exactly the pound, foot, poundal and foot poundal. Two parallel systems are no more difficult to learn than one—perhaps less difficult. Why should the student be led to believe that in an English system one must form different units in a different manner? . . . . It might seriously be claimed that the poundal made things too easy. If this simple system makes physics too difficult for students, let them elect other subjects."

I entirely agree that two (and indeed three or four) parallel and rational systems of units are no more difficult to learn than one—perhaps less difficult. Using the simplest possible defining equation, $F=ma$, there are four rational parallel systems of units:—

	Mass	Length	Time	Force
C.G.S.	Gram	Centimetre	Second	Dyne
F.P.S.	Pound	Foot	Second	Poundal
British	Slug	Foot	Second	Pound-weight
M.K.S.	Kilogram	Metre	Second	Newton

Why foment confusion by the introduction of a formula $F/w=a/g$ for the propagation of chaotic units?

REVIEWS OF BOOKS

European Science, by H. Stafford Hatfield. Pp. xiii + 152. (Cambridge: The Basic English Publishing Company, 1939.) 3*s*. 6*d*.

This book has been written in basic English. As a medium for the interchange of ordinary thoughts and of limited ideas this language doubtless is of value. As a vehicle of scientific thought and ideas, however, it leaves much to be desired. For someone just learning to speak English, provided they learn the right 850 words (and that, with a little foresight, should not be difficult), this book will be of great value and interest. An English scientific reader will, however, probably be a trifle disappointed with the substance of the book. There are inconsistencies and statements which, from a scientific point of view, are inaccurate or incorrect. Thus, for example, it is stated: " The word ' weight ' has, even in science, kept its sense of Newton's ' mass ' ; and some of the best writers today are using it where mass would have been used some years back, for example, the ' weight of an electron '." ! Since *mass* appears to be allowed in basic—the word is used elsewhere in the book—it seems a pity, in a way, that the author has not catered for those of us who are unacquainted with the " best " writers. The statement that " Heat and light were now taken in " is, even in its context, rather surprising at first sight, and one is immediately brought face to face with the fact that basic English is not an ideal medium for scientific expression.

Another example of seeming inconsistency is an historical one. " Unhappily, Darwin's work was not straight away tested by experiment ; it was attacked by the Church . . . . While this was going on, one man, Mendel, was doing some simple experiments in his garden which were a test of the idea on which Darwin's theory was based ". The fact that Mendel was a unit of that Church, later abbot of the abbey in the garden of which these experiments were carried out, does not seem to have been appreciated.

W. B. M.

A Laboratory Manual of Electricity and Magnetism, by Leonard B. Loeb. Pp. 121, with experimental sheets. (California: Stanford University Press; London: Sir Humphrey Milford, Oxford University Press, 1941.) 22*s*. 6*d*. net.

This book is a revised edition of a practical physics manual of electricity and magnetism for junior classes at the University of California, and is intended to supplement the author's *Fundamentals of Electricity and Magnetism*. The latest volume of the author is noteworthy for his unique method of instructing students, the salient features of which are clearly set out in the preface. In particular, Prof. Loeb stresses the importance of a student having at least some acquaintance with the relevant theory and basic principles of an experiment before he attempts to carry it out in the laboratory. Following this concept, the author, after titling an experiment and stating the objective to be achieved, gives an elementary description of its underlying theory. This procedure is, in the opinion of the writer, to be commended, especially in the case of large classes where it is impossible to arrange for all students to be performing experiments based on theory received in previous lectures. However, one feels that it is a somewhat retrograde step to include, in the manual, perspective line diagrams of apparatus set up in electrical circuits.

In order to avoid students spending what the author considers an unnecessary length of time in writing-up laboratory records, there is incorporated at the end of each experiment a schematic table for the insertion of the quantitative results obtained. It is unfortunate from the students' point of view that the cost of the book is somewhat higher than it might be owing to the needless repetition (repeated four times in all) of these tables in the last hundred pages.

Twelve experiments are described in all, and they emphasize such fundamental concepts as magnetic and electric fields, electric current, potential difference, self-induction, etc. The terminology is in places a little confusing: for instance, on page 66 T_0 is used to signify the torsional coefficient of a galvanometer suspension, while T indicates its period ; again, English readers may be misled by the author's definition of the *Figure of Merit* of a galvanometer. The book is remarkably free from errors, except for one, of small consequence, at the top of figure 4, where F_M should read F_H.

R. W. B. S.

The Behaviour of Slow Electrons in Gases, by R. H. HEALEY and J. W. REED. Pp. vii + 169. (Sydney: Amalgamated Wireless (Australia), Ltd., 1941; agents for Great Britain: Iliffe & Sons, Ltd.) 20*s*.

A large number of original papers have been published during the past twenty-five years dealing with the study of the motion of slow electrons in gases by the diffusion methods originated by Townsend in Oxford and developed by Bailey in Sydney. Before the publication of this attractively produced book, which deals mainly with the results of these researches, both theoretical and practical, it would have been a difficult task for an enquirer to get a general view of the progress achieved.

An account of the fundamentals of electron motion in gases and of the early experimental work is given in Chapter 1. Chapters 2 and 3 are concerned with the formation of negative ions and with the velocity distribution of slow electrons in gases, respectively. Chapter 4 gives an account of results obtained by diffusion methods, while in Chapter 5 other methods of investigation are discussed and the results compared with those given by the diffusion method. There is a particularly full account of the formation of negative ions by attachment. Chapter 6—the last—deals with practical applications of the subject matter, following preliminary theoretical discussions of ionization by collision and the motion of electrons in gases under the influence of electric and magnetic fields. The effect of thunderstorms on the ionosphere, the possibility of production of artificial aurorae and the interaction of radio waves in the ionosphere, are all treated from both theoretical and practical aspects. The conclusion that it would be possible to produce artificial aurorae with the expenditure of only about 100 kw. of power (by using a suitable array of 4050 half-wave aerials transmitting at the local gyro-frequency) is most stimulating to the imagination.

Derivations of important formulae and very full experimental data are given throughout the text. References to individual papers are given conveniently in footnotes, as well as in the bibliography at the end of the book. It is claimed that this bibliography—relating to the whole field—is comprehensive.

The authors' aim to achieve completeness of discussion will make the book particularly useful for reference purposes ; the last chapter will certainly be of great interest to all concerned with the propagation of radio waves.

W. A. L.

Endeavour, Volume 1, Number 1, edited by E. J. HOLMYARD. Pp. 48. (Published by Imperial Chemical Industries, London, S.W. 1, January 1942.) 5*s*.

This first number of a new quarterly journal shows that it does not intend to trespass on the preserves of any existing publication. It is to survey scientific achievement, particularly British achievement, and to bring the results of this survey to as wide a world-audience as possible. For this reason it will be issued in French, German and Spanish editions, as well as in English.

The importance of good presentation and typography in a journal which may to some extent be judged as a sample of British workmanship has not been overlooked. Even more important than format is the choice of articles, and here the editor has been both discriminating and fortunate. The Astronomer Royal writes on the Sun's distance, now known to within 1 part in 10,000, Fairbrother on the Cyclotron, Newbury on the life and work of Glover, and Jacks on Soil Conservation. There are also biological articles by Waddington, Bacharach and Yonge, as well as general articles on Science in Britain and in the U.S.S.R. respectively, the former being beautifully illustrated with portraits of pioneers from Newton to Sir William Bragg.

In wishing the Journal the good luck which it merits, it would be unfair to the Directors of Imperial Chemical Industries not to express the gratitude of British men of science to their public-spirited action in launching the new venture. J. H. A.

Thermochemical Calculations, by RALPH R. WENNER. Pp. xii + 384. (New York and London: McGraw-Hill Book Co., Inc., 1941.) 28*s*.

" **Problem 33.** *The free energy of Gaseous Formaldehyde.*—Calculate the ΔF° of formation of HCHO(*g*) from the following data:—Heat of combustion = 134,100 cal./mole. Fundamental vibration frequencies = 1,165, 1,278, 1,503, 1,750, 2,780, 2,875 cm.$^{-1}$

Interatomic distances: C=O distance = 1·21 A.
C—H distance = 1·09 A.
H—C—H ∠ = 120°.

" **Problem 34.** *Electrolytic reduction of Nitrobenzene to Aniline.*—Estimate the kilowatt-hours required to produce 1 ton of aniline by cathodic reduction of nitrobenzene dissolved in strong acid. Assume a current efficiency of 80 per cent. and make reasonable assumptions as to over-voltages and cell resistance."

These two consecutive items from a list of sixty exercises given at the end of the book will give a clearer idea of what the student should learn from it than would a long description. In one sense it is an introduction to chemical engineering, but the outlook is far more theoretical than in any other book on this subject known to the reviewer. In another sense, it is a sequel to Lewis and Randall's standard text-book on *Thermodynamics*, but with a more practical outlook.

As a training for the reader who, in the author's words, has " already been exposed to some formal exposition of the principles of thermodynamics " it should be invaluable. He is taught to seek auxiliary data in the literature, to estimate missing data by logical comparison with data for similar compounds, to compute specific heats of gases from spectroscopic data and to estimate compressibility data by making up an equation of state from the critical constants. In short, he is introduced to and encouraged to practise those tricks which he will have to use in later life if he is ever to make advances or to depart from established practice.

The last nine chapters deal with solubilities, titration and similar physico-chemical matters, chemical reactions, flame temperatures, converter design, gas-absorption plant, air conditioning, metallurgy, gas liquefaction, and fluid flow, each treated thermodynamically. The preceding eight chapters are provided as introductory to them, and deal with pure thermodynamics, and with methods of computing specific heats, enthalpies, Lewis free energy, and equilibrium constants. Lest the remark above, as to the calculation of specific heats from spectroscopic data, should be misleading, it should be added that the rotational and vibrational heat capacities of gases are also considered, and that solids are not overlooked.

The student brought up on these lines certainly starts his career with many advantages over his colleagues of the preceding generation. J. H. A.

The Cathode Ray Tube and its Applications, by G. Parr. Pp. viii + 180. (London: Chapman and Hall, Ltd., 1941.) 13*s*. 6*d*.

This well-produced second edition of *The Low Voltage Cathode Ray Tube*, by the editor of *Electronic Engineering*, gives a clear, well-informed and eminently readable account of the fundamental processes involved in the working of the cathode-ray oscillograph and of its many applications. In the preface, reference is made to the helpful criticism received from readers of the original edition of the book—and notably that from Mr. R. A. Watson-Watt. Although the author regrets that he could not describe some of the latest and best developments in cathode-ray technique, the book in its present compact and lucid form will undoubtedly prove most valuable to all readers who wish to obtain a general view of the field of cathode-ray oscillography. Detailed information which has been published on any particular branch of the subject will readily be obtained with the aid of the bibliography, which contains 737 references.

The range of the subject matter is so wide that the general reader might hesitate to voyage on such an ocean of specialized knowledge. However, the skill with which he is piloted here from the history, construction, operation and performance of cathode-ray tubes (chapter 1 and 2) through the intricacies of Lissajous' figures and time-base circuits (chapters 3, 4 and 5) to the diverse uses of the tube in radio engineering (chapter 6), in industry, and in many other applications, as, for example, in electromedical practice in studying heart-beats and brain-cell activity (chapter 7) to the final chapter on television reproduction, will amply repay his embarkation.

There are 80 figures, chiefly in the form of good line diagrams, and an Appendix on photography.

The book will serve as a most useful introduction to the subject for those who are preparing for future specialized training. W. A. L.

THE PROCEEDINGS OF
THE PHYSICAL SOCIETY

VOL. 54, PART 3 1 May 1942 No. 303

WILLIAM DAVID COOLIDGE

EIGHTEENTH DUDDELL MEDALLIST

THE eighteenth Duddell Medal was presented to Dr. W. D. Coolidge by His Excellency the British Ambassador, the Right Honourable Viscount Halifax, at a dinner given on 1 May during the meeting of the American Physical Society at Baltimore.

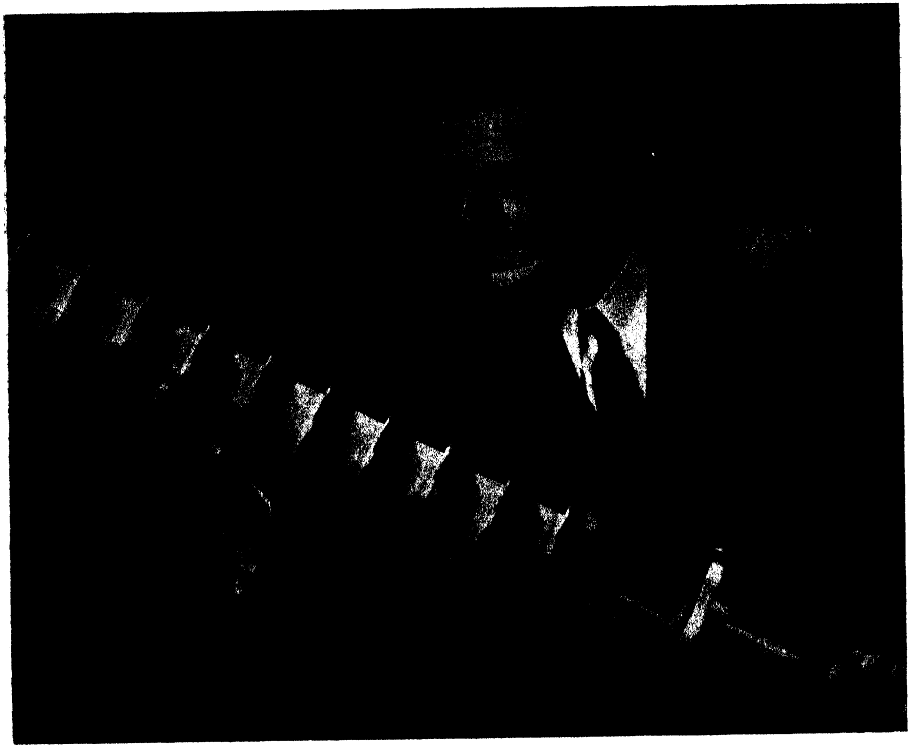
Coolidge is best known to physicists for his x-ray tube, and for his share in the development of the tungsten filament, either of which falls within the category of "development of instruments or materials", which it is the prime purpose of the Duddell Medal to recognize.

William David Coolidge, this year's medallist, was born in 1873 and passed his early life on a farm in Massachusetts; he graduated at the Massachusetts Institute of Technology in 1896, at a time when the first announcement of the discovery of x rays was exciting interest everywhere. He remained at the M.I.T., except for a short period of research at Leipzig, until 1905, when he went to the research laboratories of the General Electric Company, Schenectady, where he has remained ever since. He became Assistant Director of the Laboratories in 1908, Associate Director in 1932, and finally a vice-president of the company and Director of Research since 1940.

His work on tungsten, as is well known, led to the possibility of producing this usually brittle metal in ductile form, and so made it possible to use it in electric lamps, where its high melting point permits the filament to run at a very high temperature, so as to obtain great luminous efficiency. The value of the work to physicists, however, has been not only in the possibilities which it opened up in the field of instrument construction, but because it focused attention on some of the problems of the solid state. In this way, his work has contributed indirectly to this fundamental problem of theoretical physics.

His hot-cathode x-ray tube was described 30 years ago; it differed from earlier tubes in that the current and voltage could be controlled independently. In his tube, which is very highly evacuated, variation of the current through the filament alters the electron emission from the cathode, and thus gives control of the tube current, and with it of the intensity of the radiation. The voltage between anode and cathode can be varied independently of this, thus enabling radiation of varying "hardness" to be produced. Later, Coolidge developed oil-immersed self-contained x-ray units, which have been of great value in medical and industrial work. A further development was announced in 1924, when he described a tube for high voltages. This tube was sub-divided, and accelerating voltages applied to the various sections. In the accompanying photograph, the designer is seen holding a tube designed for one million volts. Such tubes have enabled x rays of great penetrating power to be applied medically, and have enabled deep therapy to be used under closely controlled conditions. They have also been applied industrially in radiographic examination of materials.

Another outcome of Coolidge's work has been the study of the cathode rays outside the tube, a development which aroused considerable interest both in the technical and the lay press a few years ago. The cathode-ray particles emerge from the tube through an exceedingly thin window of metal foil, and their track in air is brilliantly marked. In living tissue, their penetration with the voltages available at first was very small, but with higher exciting voltages, rays of great penetrating power may be anticipated. These promise to be of therapeutic value in certain maladies.



Coolidge inspecting his million-volt x-ray tube.

METEOROLOGICAL RESEARCH

§ 1. METEOROLOGICAL RESEARCH COMMITTEE OF THE AIR MINISTRY

The importance of meteorology to aviation has necessitated an extension of the State meteorological service far beyond the modest dreams of meteorologists of the last generation. At the present time meteorological forecasting is a vital element in the national war effort, especially in the conduct of the offensive ; but the return of peace and the renewed expansion of civil aviation will maintain the need for weather prediction, which also has great importance for agriculture and many other branches of the national economy.

The science and art of weather prediction have made great progress during the present century, but it remains uncertain whether even now all the fundamental factors that determine the weather changes have been discovered, and also whether the programme of observation yet includes all the elements necessary as a foundation for prediction. The problems both of observation and of theory are extremely complicated. It is therefore desirable that, along with the day-to-day application of the knowledge already available, proportionate efforts should be made to improve our knowledge and our methods.

For this reason the Secretary of State for Air has recently appointed a Meteorological Research Committee to advise and assist in the carrying out of meteorological investigations. The Chairman of this Committee is Professor S. Chapman, and the members are Professor D. Brunt, Dr. G. M. B. Dobson, Professor G. I. Taylor, the Director of the Meteorological Office, the Director of Scientific Research (Ministry of Aircraft Production), the Director of the Naval Meteorological Service, and representatives of the Air Staff and Civil Aviation.

At present the Committee will naturally be concerned chiefly with problems directly concerned with the war effort, and such work must for the time being remain secret, though any incidental results not likely to be of service to the enemy may be published.

The Committee would welcome contact and co-operation with university departments or other institutions engaged on work that bears on meteorology ; as a result of enquiries in the United Kingdom, information concerning such work and offers of assistance have been received from several universities. Correspondence* is invited from any other university or research institutions which can assist in this work.

§ 2. METEOROLOGICAL RESEARCH IN THE ROYAL SOCIETY

The formation of a Meteorological Research Committee by the Air Ministry, reported above, is an important and very welcome advance in the organization of British meteorological research. The Air Ministry Committee will at present confine its work chiefly to those problems which have an immediate practical application and are likely to be solved in a fairly short time. The Royal Society has been invited by the Air Ministry to co-operate with the Meteorological Research Committee by undertaking research on certain aspects of meteorology which, though of fundamental importance for the advance of the subject, may not have an immediate practical application.

The Council of the Royal Society have agreed to this request, and have entrusted the immediate responsibility for the work to the Gassiot Committee. This Committee, first constituted in 1871, was originally appointed to administer the Gassiot and other Trust Funds applicable to the maintenance of certain British meteorological and magnetic observatories, and to make recommendations as to their work. The terms of reference have now been enlarged to include the supervision of fundamental meteorological research such as has been asked for by the Air Ministry. The personnel of the Committee is appropriate for this purpose ; the chairman is Dr. G. M. B. Dobson, and the members are Sir Edward Appleton, Professors D. Brunt, S. Chapman, and A. C. G. Egerton, Sir Henry Lyons, Sir George Simpson, Professor G. I. Taylor, Sir Gilbert Walker, the Astronomer Royal, the President of the Royal Astronomical Society, the Director of the Meteorological Office, and, *ex officio*, the Treasurer of the Royal Society (Professor T. R. Merton) and the Secretary of the Royal Society (Professor A. C. G. Egerton).

*** Addressed to the Secretary, Meteorological Research Committee, Meteorological Office, Air Ministry, Kingsway, London, W.C. 2.**

These arrangements seem eminently suitable, because pure fundamental research which may have to be continued over a long period, and which requires the co-operation of many independent workers in university departments and laboratories, can probably be organized more easily by the Royal Society than by a Government Department, however much the latter may be in sympathy with the object of the researches. The Royal Society has already shown its willingness to further such work by appointing Dr. E. Glückauf to a Mackinnon Studentship, to enable him to carry out meteorological research under the general guidance of the Gassiot Committee.

The general research in which the Air Ministry has sought the co-operation of the Royal Society is that of radiation in the atmosphere, with special reference to radiation equilibrium conditions in the stratosphere. It is fairly generally accepted that the average temperature of the stratosphere—as of the warm region above it—is determined by an equilibrium between the energy absorbed, from both the downward solar radiation and the upward terrestrial radiation, and the energy radiated by the air in the stratosphere. Observations have shown that there are large variations of temperature in the stratosphere both with time and with place, but we know little as to why such variations occur. Few meteorologists, for instance, would care to be dogmatic as to why the stratosphere is some 50° c. *colder* in equatorial regions than in polar regions.

The principal gases of the atmosphere are mostly very transparent to radiation, and the radiative conditions are largely governed by such minor but polyatomic constituents as water vapour, carbon dioxide, ozone, and possibly others. To understand the radiation equilibrium in the atmosphere it is necessary to know (*a*) what gases are present which have important absorption bands for radiation; (*b*) in what proportions thay are present at different levels ; (*c*) how gases such as ozone are formed and decomposed ; (*d*) the absorption coefficients of these gases for different wavelengths, and whether the absorption and emission spectra have a continuous or fine line structure. Finally, when the above facts are determined, it will be necessary to calculate the radiation equilibrium temperatures at different levels and to determine how rapidly the temperature will change when the conditions alter.

The Gassiot Committee has been fortunate in enlisting expert help in particular branches of the subjects involved, and has formed three sub-committees to deal with different aspects of the above question. It is hoped in this way that the available relevant knowledge will be collected, and that those questions which require more experimental work may be defined, so that arrangements can be made for the work to be done, so far as this should prove possible under present conditions.

Dr. Dobson is the chairman of sub-committee A, which will deal with the chemical analysis and the observation of the amounts of the polyatomic constituents of the air at different levels ; other members of this sub-committee are Professors H. Dingle and F. A. Paneth, and Dr. E. Glückauf.

Sub-committee B will deal with the photo-physics and photo-chemistry of the atmosphere, from the highest levels downwards, and, in particular, with the changing equilibrium of ozone and any other variable polyatomic constituents ; the chairman of this sub-committee is Professor H. W. Massey, and the members are Professor K. G. Emeléus, Dr. W. C. Price, and Dr. J. Sayers.

Sub-committee C will consider the radiation balance in the atmosphere ; its chairman is Professor S. Chapman, and its members are Professor D. Brunt, Dr. G. S. Callendar, Dr. T. G. Cowling, Dr. A. R. Meetham, Sir George Simpson, and Dr. G. B. B. M. Sutherland.

There will be few who will not agree that the first meteorological problem thus selected for attack is a very important one, and still fewer who will dispute its difficulty. It is hoped that the scientific effort which the Royal Society can bring to bear upon this work will lead to substantial progress.

The Gassiot Committee is aware that at laboratories and institutions in many parts of the world work has been undertaken that bears on these problems, and the Committee would welcome co-operation in cases where the continuance of such work is possible in present circumstances.

THE PHYSICS OF BLOOD-PRESSURE MEASUREMENT

By DAVID S. EVANS* AND K. MENDELSSOHN,
Clarendon Laboratory, Oxford

MS. received 17 *December* 1941

ABSTRACT. The physical basis of blood-pressure measurement has been formulated. An arrangement for recording the essential features of the pulse curve is described, and the relation between the recorded values and the pressure changes in the artery discussed. The results have been confirmed by clinical observations.

§ 1. INTRODUCTION

In wartime, the determination of arterial blood-pressure in casualties is a matter of the greatest importance, since it is the only means for the early detection of the serious condition known as shock. It seems, therefore, worth while to discuss in detail the physical basis of such determinations and to explore the possibilities of the development of existing methods in order to improve their accuracy and reproducibility.

The action of the heart produces in the arteries a pressure variation which is more or less periodic. Medical practice is accustomed to define three parameters in connection with this variation. They are :—

(i) The frequency, or *pulse rate.*

(ii) The maximum pressure attained in the artery during the pulsation. This is the *systolic pressure*, p_s.

(iii) The minimum, or *diastolic pressure*, p_d.

The quantity, $p_s - p_d$, is called the *pulse pressure*, which we have denoted by p_0.

Although the description of the pulse curve given by these parameters is rather rough, they are the only quantities which have hitherto been readily determinable in practice. Typical values for a healthy individual are:— Systolic pressure: 100 to 150 mm. of mercury, with pulse pressure about one-third of systolic. Pulse rate: 60 to 120 per minute. In cases of shock the systolic pressure may be as low as 40 to 60 mm.

For practical reasons it is impossible to tap the artery directly in order to measure the pressure. The usual arrangement for measurement consists of a rubber cuff, A (see figure 1), which can be inflated with air to any desired pressure. It is secured in position around the limb l by means of an inextensible bandage. The air pressure in the cuff is read on the manometer m_1. For our purposes, the tissues, t, of the limb may be regarded as an incompressible jelly, which transmits the external pressure applied by the cuff. Thus, when the pressure in A is raised to a value equal to the systolic pressure in the artery a, the latter collapses completely throughout the pulse cycle, and cuts off the flow of blood. The systolic pressure can, therefore, be determined with the aid of any device which will indicate the cessation of blood-flow past A, the pressure at which this happens being read on m_1.

---

* By kind permission of the Board of Visitors of the University Observatory, Oxford.

The usual method is to employ a stethoscope for detection, and systolic pressure is indicated by the first appearance of pulse sounds as the external pressure is lowered. The diastolic pressure is read in a similar manner, the indication in this case being that the arterial sounds grow muffled or disappear entirely when the pressure in A falls to a certain value.

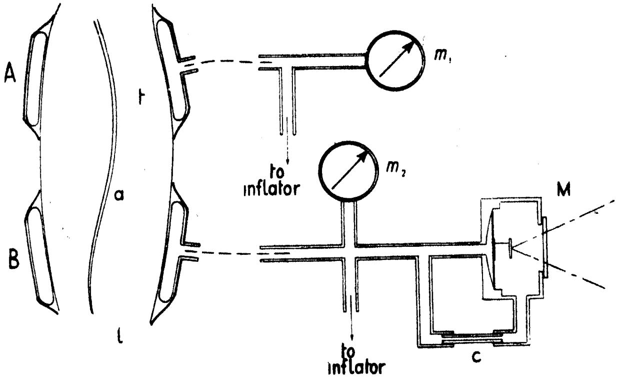


Figure 1.

The reason for this change in the sounds is unknown, but we think it may be due to a change from turbulent to laminar flow in the artery. There exists no theoretical reason for identifying the pressure at which this change in sound occurs with the diastolic pressure as defined. Besides being theoretically unsound, the auscultatory method has the disadvantage of yielding values on the same patient which vary widely with individual measuring technique, and it is difficult to carry out in an operating theatre or casualty clearing station, where conditions are not conducive to the accurate detection of faint sounds in a stethoscope.

Amplification of the pulse sounds, or their conversion into a visual record by electrical methods, is ruled out, because simplicity and cheapness are essential factors in the solution of the practical problem of producing a recording instrument.

§ 2. EXPERIMENTAL ARRANGEMENT

The detector chosen after a number of exploratory trials was a second cuff, B, which was placed on the limb below A.*

---

* The use of a second cuff, or the combination of two cuffs, is by no means new, but the various modes of operation hitherto employed are fundamentally different from ours.

This recording cuff, B (figure 1), is exactly similar to A, and consists of a rubber bag and an inextensible bandage to strap it around the limb. It can be inflated to any desired pressure, which is indicated on the manometer m_2. Pulsations in the artery are transmitted to this cuff, and are measured by means of a highly sensitive membrane manometer, M, whose deflections are recorded by means of a light beam and scale. This manometer consists of two chambers, of which the smaller is directly connected to the cuff, while the connection with the larger is made through a capillary, c. This arrangement serves to record only the pressure changes of short period. If the dimensions of the capillary and manometer are correctly chosen (see § 4), it records the patient's pulse faithfully, but is undisturbed by the very much larger, but gradual, volume changes due to the variation of temperature or fluid content of the tissues.

§ 3. THE DETERMINATION OF p_s, p_d AND p_0

We shall now discuss in detail the relation between the record provided by the instrument and the true values of systolic, diastolic and pulse pressure in the artery.

We consider first the amplitude of the pressure pulsations in the inflated cuff B, as indicated by M, in relation to the mean pressure in B. The upper cuff, A, is left deflated. We define the following quantities:—

p, the pressure within the artery at any instant.
P, the mean pressure of air in the cuff B.
v, the volume at any instant of that part of the artery covered by B.
V, the total volume of B and the recording circuit.
M, the mass of air contained by V.

Since the cuff is restrained by an inextensible bandage, any increase in the volume of the artery must be compensated by an equal decrease in the volume of the cuff, or

$$v + V = \text{constant}. \qquad \ldots\ldots(1)$$

Further,

$$PV = kM, \qquad \ldots\ldots(2)$$

where k is a constant for the system under consideration. If, during the pulsation, the volume of the artery increases by an amount Δv, there will be a corresponding increase of pressure ΔP in the cuff, given by

$$\Delta P = P.\Delta v/V, \qquad \ldots\ldots(3)$$

because Δv is very small in comparison with V.

The artery is an elastic tube of negligible rigidity, so that the arterial volume v is given by

$$\left.\begin{aligned} v &= v_0 + \lambda(p-P) \text{ when } p > P \\ &= 0 \qquad\qquad\quad \text{ when } p < P, \end{aligned}\right\} \qquad \ldots\ldots(4)$$

where v_0 and λ are constants depending on the individual patient.

The changes in arterial volume caused by the pulse are of three different types corresponding to the three conditions:—

(i) External pressure less than the diastolic pressure ($P<p_d$),

(ii) External pressure intermediate between systolic and diastolic pressures ($p_d<P<p_s$),

(iii) External pressure greater than systolic ($p_s<P$).

These three cases correspond respectively to the regions I, II and III in figure 2, which illustrates the relation of the amplitude, ΔP to P, the mean pressure in the cuff B.

In case (i), the artery merely pulsates freely between the extreme values $v_0+\lambda(p_s-P)$ and $v_0+\lambda(p_d-P)$ given by equation (4), so that

$$\Delta v=\lambda(p_s-p_d)=\lambda p_0, \qquad \ldots\ldots(5)$$

and hence in this case

$$\Delta P=\frac{P}{V}\cdot\lambda p_0, \qquad \ldots\ldots(6)$$

i.e., the amplitude of oscillation increases linearly with increasing external pressure, and tends to the value $\lambda p_0 p_d/V$ as P approaches diastolic pressure. This is represented by the straight line shown in region I of the figure.

In case (ii) the change in volume of the artery during the pulse cycle is from $v_0+\lambda(p_s-P)$ at systolic pressure to zero volume at diastolic, so that now

$$\Delta P=\frac{P}{V}(v_0+\lambda p_s-\lambda P). \qquad \ldots\ldots(7)$$

This relation represents the amplitude of the pressure oscillations in the cuff between the values $P=p_d$ and $P=p_s$, subject to certain reservations which will presently be discussed.

The curve of ΔP against P in region II is a parabola, which attains its maximum at

$$P=\frac{v_0}{2\lambda}+\frac{p_s}{2}, \qquad \ldots\ldots(8)$$

which is greater than p_s if $v_0>\lambda p_s$, and this appears to be the case, since on the whole the variations of arterial volume in an artery not subject to external interference are rather small compared with its total volume. Equation (7) shows that as P approaches p_d from above, ΔP tends to the value $\frac{p_d}{V}(v_0+\lambda p_0)$, while we have seen that if diastolic pressure is approached from below, ΔP attains the value $\lambda p_0 p_d/V$. Thus the amplitude ΔP of the pressure oscillations in B passes through a discontinuity at a value of P equal to the true diastolic pressure in the artery. The change in amplitude at this discontinuity is given by $v_0 p_d/V$. The occurrence of this discontinuity provides us with an ideal method for the determination of accurate values of the diastolic pressure, and has been successfully employed for this purpose in our clinical work.

We now consider case (iii). It might be expected that there would be no pulsations in B when the external pressure was greater than systolic, since the artery should then be completely collapsed throughout the pulse cycle. This, however, is not the case, since the pressure exerted by the cuff decreases in the neighbourhood of its edges. It follows that the short section of artery covered by the upper edge of B fills with blood at the peak of each pulse wave, and causes pressure oscillations in B even when $P > p_s$. The amplitude does not therefore fall abruptly to zero as P crosses p_s, but dies away gradually.

A second disturbing effect is of importance in the neighbourhood of systolic pressure. The artery below B remains empty during the greater part of the pulse cycle, and the part of the cycle during which the artery is open is so short that the flow of blood at each pulse beat is insufficient to fill completely the section of artery covered by B.

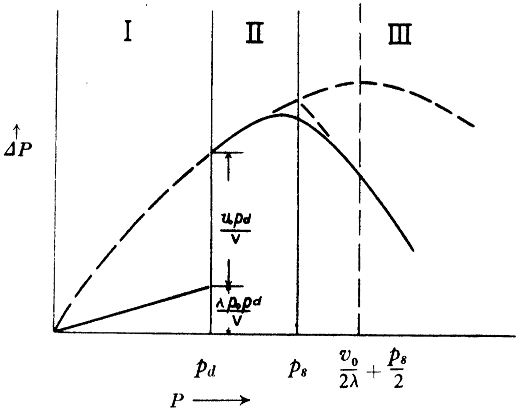


Figure 2.

The amplitude, therefore, decreases below the value given by the parabola even before p_s is reached (see figure 2). The situation is not well adapted to mathematical description, and p_s cannot be determined with any accuracy using only a single cuff.

For the determination of the systolic pressure we have therefore employed the cuff A as well as B. The latter is inflated to a moderate pressure below diastolic, and then serves as a detector for blood flowing past A. If the pressure in A is allowed to fall through the systolic value, pulsations will start suddenly in B as soon as blood begins to flow past A, thus allowing systolic pressure to be determined. With this arrangement edge effects are confined to A, and the determination is a critical one.* By observing the sudden changes in the recorded pulsations in relation to the pressures in cuffs A and B we are, therefore, able accurately and unambiguously to determine systolic and diastolic pressure, the two parameters on which current medical practice relies. It appears, however, from an analysis of the (ΔP, P) curve that our arrangement will be capable of recording further valuable information concerning the patient's circulation. Reference to equation (6) shows that when the external pressure is less than p_d the amplitude of oscillation in B is proportional to p_0. Now p_0 is the pulse pressure, which gives a direct indication of the quantity of blood flowing through the limb at each heart beat. Medical opinion attaches an increasing importance to this quantity, since a drop in pulse pressure seems to be the earliest and most reliable indication of the onset of shock. Previously p_0 could only be deduced from separate determinations of p_s and p_d, a process involving a double error and a considerable expenditure of time and attention. Our arrangement enables the observer to form a continuous estimate of p_0 itself at each moment, from the amplitude of the oscillations recorded by the manometer when B is inflated to a constant pressure below diastolic.

---

* Cuff A must, of course, be wide enough to exert an even pressure over a considerable length of the artery.

§4. THE RECORDING OF THE PULSE CURVE

So far we have been concerned only with the amplitude of the pressure oscillations in B. We now consider their detailed form as well. We take the case when the external pressure is less than p_d, and consider an instant when the arterial pressure is p. Then the arterial volume is given by $v_0+\lambda(p-P)$; at diastolic pressure it is given by $v_0+\lambda(p_d-P)$, so that the change of volume is $\lambda(p-p_d)$, and hence the pressure change in the cuff is, by equation (3),

$$\Delta P=\frac{P}{V}\cdot\lambda(p-p_d). \qquad \ldots\ldots(9)$$

Thus, as the arterial pressure p alters with time, the pressure variations in the cuff trace out a curve which is similar to the variation of pressure within the artery. This relation at once suggests that it should be possible to produce an instrument to record accurately all the pressure variations within the artery throughout the pulse cycle, and immediately raises the problem of the faithfulness of the record provided by our apparatus. Distortion of this record arises both from the action of the capillary and from the natural period and damping of the moving parts of the manometer.

The volumes of the two chambers of the manometer M are denoted by V_1 and V_2, and the pressures in them are P_1 and P_2. The division between the two chambers (see figure 1) is a metal membrane, and they are connected by a capillary of length l and diameter d, through which excesses of pressure in the one chamber leak slowly into the other, until a mean pressure P_0 is attained. The difference P_1-P_2 is very small, and never exceeds one or two mm. of water.

Then we have

$$P_1V_1+P_2V_2=P_0(V_1+V_2). \qquad \ldots\ldots(10)$$

If P_1 is greater than P_2, air will flow from one chamber to the other at a rate given by $\kappa(P_1-P_2)$ c.c. per second (assuming laminar flow), where

$$\kappa=\pi d^4/128\eta l \qquad \ldots\ldots(11)$$

and η is the viscosity of air ($=1{\cdot}8\times10^{-4}$ c.g.s. units). This volume is measured at the mean pressure, which we may take as P_0, since P_1-P_2 is very small. Then we have the equations

$$\frac{dP_1}{P_0}=-\frac{\kappa}{V_1}(P_1-P_2)dt;\quad \frac{dP_2}{P_0}=\frac{\kappa}{V_2}(P_1-P_2)dt,\qquad\ldots\ldots(12)$$

of which equation (10) is an integral.

Elimination of P_2 from the first of (12) gives

$$dP_1=-\frac{2\kappa}{V_0}(P_1-P_0)dt,\qquad\ldots\ldots(13)$$

where

$$\frac{2}{V_0}=\frac{1}{V_1}+\frac{1}{V_2},$$

and so

$$\left.\begin{aligned}P_1-P_0&=(P_1^0-P_0)e^{-\phi t},\\ P_2-P_0&=(P_2^0-P_0)e^{-\phi t},\\ \phi&=2\kappa P_0/V_0,\end{aligned}\right\}\qquad\ldots\ldots(14)$$

where P_1^0 and P_2^0 are the initial pressures.

If $F(t)$ is the pressure variation with time recorded by the manometer, and $f(t)$ the true pressure variation in the cuff, then

$$\frac{dF(t)}{dt}=\frac{df(t)}{dt}-\phi F(t),\qquad\ldots\ldots(15)$$

since the pressure decays at a rate given by $-\phi F(t)$.

Integrating (15) from $-x$ to t gives

$$\begin{aligned}F(t)&=e^{-\phi^t}\int_{-x}^{t}e^{\phi t}\frac{df(t)}{dt}dt\\ &=f(t)-\phi\int_0^\infty f(t-x)e^{-\phi x}dx,\end{aligned}\qquad\ldots\ldots(16)$$

where x is a dummy variable.

When ϕ is small (i.e. the capillary is very narrow), $F(t)$ approximates to $f(t)$, but we cannot dispense with the capillary entirely, since our highly sensitive manometer must act differentially if it is to record small fluctuations superposed on a very high mean pressure liable to slow changes. The practical limit to the smallness of ϕ is set by the necessity of preventing serious deviations of mean pressure between the two chambers. As ϕ increases, $F(t)$ tends to zero, and since ϕ is proportional to P_0, the operating pressure should be as low as possible for accurate reproduction of the pulse curve. On the other hand, equation (9) shows that the sensitivity is proportional to the cuff pressure, so that ϕ should be as small as possible, and the apparatus should be operated as near as possible to diastolic pressure.

If $V_1=V_2=10$ c.c., and $P_0=30$ mm. of mercury, then a capillary 5 cm. in length gives $\phi=21{\cdot}5$ when $d=1$ mm., $\phi=1{\cdot}34$ when $d=\frac{1}{2}$ mm. and $\phi=0{\cdot}27$ when $d=\frac{1}{3}$ mm.

In figure 3, the effect on a typical pulse-pressure curve is illustrated. The pulse rate is taken to be 1 per second, so that if the pulse rate is 120 per minute, the values of ϕ given above must be halved, i.e., more rapid pulses are recorded more accurately. We see that $\phi=1$ already gives quite a good reproduction of the curve form. If we are merely interested in the amplitude, as we are when measuring p_s, p_d or p_0, a much wider capillary can be used, since the reduction in amplitude from $\phi=1$ to $\phi=10$ is not serious. The recorded amplitude is proportional to the true amplitude, so that the analysis of the $(\Delta P, P)$ curve made in § 3 still holds.

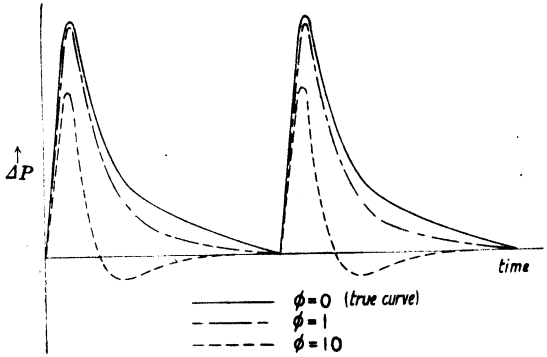


Figure 3.

Calculation of the distortion produced by the natural period and damping of the instrument shows that good reproduction can be obtained with a period equal to 1 % of the pulse period, but that a ratio of only 10 to 1 may be tolerated if the damping is considerable.

§ 5. CONCLUSION

An experimental model embodying the principles discussed in this paper has been constructed and used in extensve clinical tests. A detailed description of the apparatus and the clinical results will appear in a medical journal. These results have borne out the features of the theoretical $(\Delta P, P)$ curve (figure 2). The continuous observation of p_0 has been carried out during lengthy operations, and has proved of considerable value both to anaesthetist and surgeon. No use has been made so far of the detailed form of the pulse curve whose observation was discussed in § 4, since it was our aim to produce a recording apparatus simple and cheap enough to permit mass production. There is, however, no doubt that an accurate knowledge of the pulse curve will be of great importance in many cases, and this record could easily be obtained from our arrangement. The production of a second model fitted with a film camera is under consideration.

§ 6. ACKNOWLEDGEMENTS

In conclusion, we wish to express our thanks to Professor R. R. Macintosh, of the Department of Anaesthetics, Oxford University, who formulated the medical problem, and to the staff of his department for their interest and help in the clinical experiments. We are also indebted to the surgeons of the Radcliffe Infirmary, Oxford, who accorded us facilities in the operating theatre. Finally, our thanks are due to Mrs. V. Launchbury, whose voluntary assistance has been invaluable throughout the experimental work.

COLOUR BLINDNESS AND ITS IMPORTANCE IN RELATION TO INDUSTRY*

By F. H. G. PITT,
Developments Dept., Kodak Ltd.

MS. received 29 *January* 1942

It is common knowledge to all persons interested in colour that the colour-vision characteristics of individuals vary considerably. This variation is so great that whereas one observer will say that two colours are completely different, another may say that they are almost identical. Persons whose colour vision is different from that of the normal, and whose range of colours is in general more limited, may be broadly defined as being colour blind. While this term is rather misleading, the implication being that all colour-blind persons are blind to colour, it has been so long and so universally used in the literature on the subject as to render its replacement by a better term almost impossible. Colour blindness is almost always congenital, and as such is incurable. There is good evidence that colour blindness may be acquired, but reliable information on this aspect of the subject is very scarce, and for this reason only the inherited type will be discussed.

Much information has been collected on normal and abnormal colour-vision, and it is possible, without any reference to the numerous colour-vision theories which exist, to classify scientifically the types of colour vision of the human race into three main classes, which may be in turn further subdivided. This classification may be arranged as follows:—

Trichromats

This class contains those whose colour vision is a function of three variables, it being possible for an observer in this class to express any colour, C, uniquely by an equation of the type

$$C=\alpha R+\beta G+\gamma B,$$

---

* A paper read to the Colour Group 11 December 1941

where R, G and B may, for example, be the primaries of a colorimeter and α, β and γ the amounts required to match the colour C.

This class may be subdivided into

(*a*) Persons with normal colour-vision,
(*b*) Persons with abnormal colour-vision, called Anomalous Trichromats.

Dichromats

This class contains those whose colour vision is a function of two variables, and may be subdivided into

(*a*) Protanopes, called red-blind by Helmholtz and red-green blind by Hering;
(*b*) Deuteranopes, called green-blind by Helmholtz and red-green blind by Hering;
(*c*) Tritanopes, called blue-blind by Helmholtz and blue-yellow blind by Hering.

Von Kries introduced the Greek names, which were meant to be descriptive only.

Monochromats

This class contains those whose colour vision is a function of one variable only; cases are extremely rare. This class, together with that of Tritanopes, also rare, are not deemed numerous enough to be taken into account in a paper of this description. As both classes are generally due to disease of the eye, their defect should never go undetected.

Colour blindness is only abnormal when referred to normal vision. It is, therefore, necessary to compare the results from both types for a full understanding of the subject. As it will be shown later that anomalous trichromatism may yield varying characteristics, whereas those of the dichromats form a fairly well-defined set, comparison between the latter class and normal vision will be made first. It is obvious that this comparison is best obtained by the knowledge of how these classes of observers see certain colours. The spectrum, with its large range of pure colours of varying intensities, which can be produced under closely controlled conditions, forms an admirable test object. It is possible for the observers to measure the change of hue, the intensity, and the relative amounts of certain pre-determined primaries from one wave-length to another and thus record exactly how the spectrum appears to them. This information is of extreme importance, for on it is based—knowingly, or sometimes unknowingly—all tests used in the detection of colour blindness. It has the advantage of paying no attention to nomenclature, which undoubtedly differs markedly from observer to observer.

Hue discrimination

The measurement of hue-discrimination data offers more difficulty than that of luminosity or colour mixture. The proper conditions of observations are important, and it is essential that the instrument used should be easily controlled by the observers, as this type of measurement is of a tiring kind. Unless these conditions are observed, and unless specially suitable apparatus is used, the results tend to become a function of variables other than the actual observers. In practice, the method used is to find the just-perceptible difference between two juxtaposed fields in terms of wave-length. Results shown in figures 1 and 2 are for normal trichromatic and dichromatic observers, for a 2° field, and were all made by means of the same apparatus, the Wright Colorimeter (Wright, 1927/28), on which, incidentally, all the other measurements described were made.

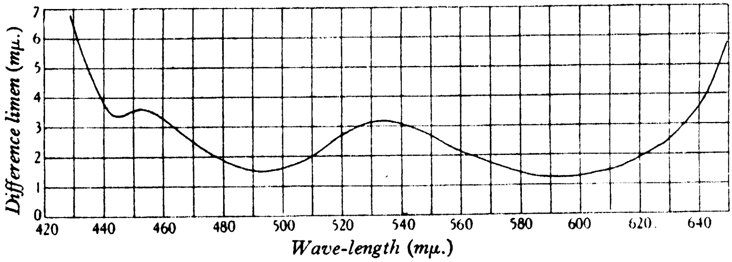


Figure 1. Hue-discrimination curve for normal observer (Wright and Pitt, 1934).

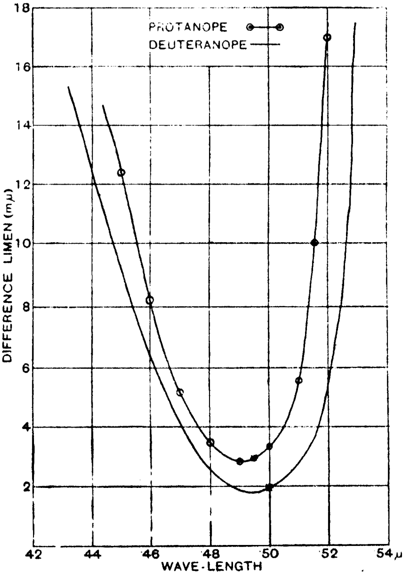


Figure 2. **Mean hue-discrimination curves from the results of 6 protanopes and 6 deuteranopes (mean position of neutral points marked with asterisks) (Pitt, 1935).**

The differences between the trichromatic and dichromatic observers are probably the most outstanding of all the differences. The curves are in themselves self-explanatory, and from them it can be calculated that the average deuteranope can see about 30 just-distinguishable spectral colours and the average protanope about 20 just-distinguishable spectral colours, while the normal trichromat can distinguish about 150 spectral colours. This comparison between trichromat and dichromat does not, however, present the full story. Colour for dichromats is a function of two variables only. As one coefficient increases the other correspondingly decreases, and it therefore follows that all colours possible to dichromatic vision, whether spectral or otherwise to the dichromats, can be seen by dichromats in the spectrum. It has been estimated, as a rough approximation, that the normal trichromat can distinguish about 5000 different colours when a 2° matching field is used. In comparison, the dichromat lives in a very limited world of colour (see Konig, 1903; Steinder, 1906; Jones, 1917; Laurens and Hamilton, 1923; Wright and Pitt, 1934; Hecht and Schlaer, 1936).

Luminosity

The standard luminosity or visibility curve for an equal-energy spectrum, as proposed by Gibson and Tyndall (1923), was adopted by international agreement at the meeting of the International Commission of Illumination of 1924. This curve does not represent the observations of one particular person, but was deduced from the data of various researches on the subject, and really represents the average eye of 250 normal trichromatic observers. In figure 3 a normal curve is shown together with the mean luminosity curves from the results of 6 protanopes and 6 deuteranopes.

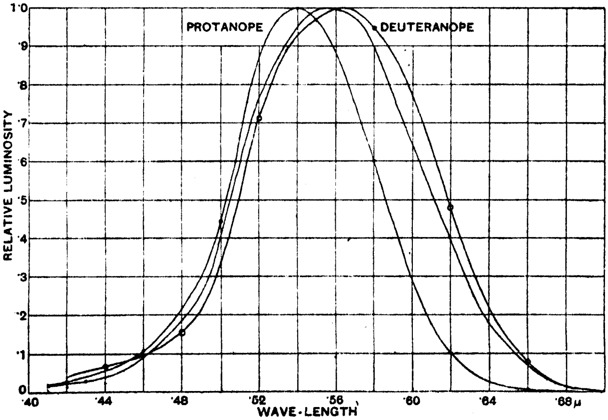


Figure 3. Mean luminosity curves (equal-energy spectrum) from the results of 6 protanopes and 6 deuteranopes; also normal curve (Pitt, 1935).

The essential differences of the dichromatic curves are apparent. The protanopic curve has a sharper maximum, the shape of the curve being narrower throughout. Probably the most important point to notice is that the ordinates of the protanopic curve are considerably smaller from $0{\cdot}56\,\mu$ to the red end of the spectrum: this is the only essential difference between the protanope and deuteranopes. From $0{\cdot}53\,\mu$ to the blue end of the spectrum the differences between all three luminosity curves are not marked, and the curves undoubtedly overlap for individuals in the three groups. The red ordinates of the deuteranope tend to be higher than those of the normal trichromat (see Pitt, 1935; Wright and Pitt, 1935; Hecht and Schlaer, 1936).

Colour mixture

For the normal observer, any colour may be represented in terms of three other colours. That is,

$$C = \alpha R + \beta G + \gamma B,$$

where C is any colour, and R, G and B are three other colours. In modern colorimetric work the sum of the coefficients α, β and γ is made to equal unity, as this specifies the quality of the colour C uniquely, although no indication of the luminosity is given.

In the same way, for the dichromatic observer, any colour may be represented in terms of two other colours. That is,

$$C = \alpha R + \beta B.$$

The data of figures 4 and 5 used the W.D.W. system of units (Wright, 1929), and can be transformed readily to the better known C.I.E. system (Guild, 1924/25; Pitt, 1935). Figure 4 is the mean curve calculated from the results of 10 normal observers. Mean curves from the results of 8 protanopes and 7 deuteranopes are shown in figure 5. The differences between figures 4 and 5 are to be

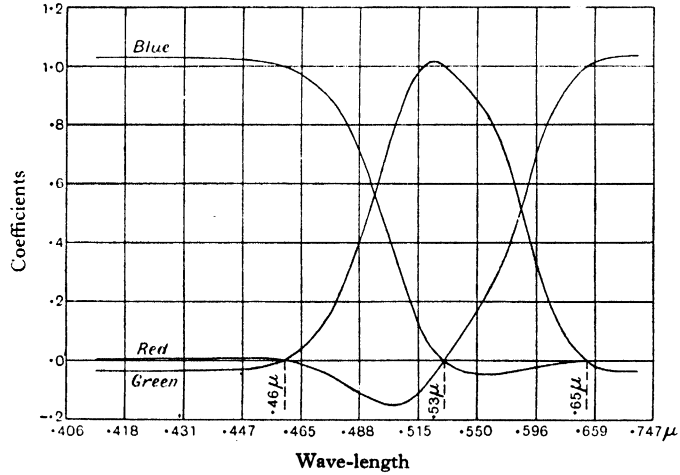


Figure 4. Mean coefficient curves from results of 10 observers (Wright, 1929).

expected from the foregoing remarks. The two mean curves shown in figure 5 show a difference from $\lambda = 0{\cdot}51\mu$ to the red end of the spectrum, but it should be pointed out that the coefficient curves cannot be used as an infallible test to show whether the dichromat in question is a protanope, as the individual curves of both classes overlap (see Konig and Dieterici, 1892; Abney, 1913; Wright, 1929; Pitt, 1935.)

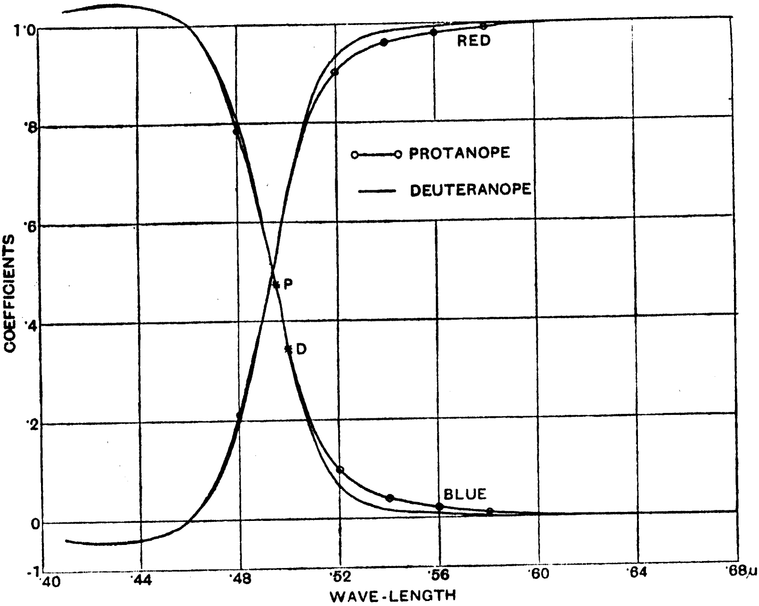


Figure 5. Mean coefficient curves from the results of 8 protanopes and 7 deuteranopes (position of neutral point marked with asterisks) (Pitt, 1935).

Anomalous trichromats

In general, anomalous trichromats differ from normal trichromats in that, although they have a colour system which is a function of three variables, the colour matches may differ from those of a normal trichromat. It is not possible to present the results of measurements made by anomalous trichromats in the same way as those of the normal trichromats or dichromats have been presented, as anomalous trichromatism ranges progressively between normal trichromatism and dichromatism (Nelson, 1938; McKeon and Wright, 1940). Figure 6, due to McKeon and Wright, shows hue-discrimination curves of 11 protanomalous observers arranged in order of increasing magnitude of defect. From these curves it can be seen that hue discrimination becomes progressively poorer in the greenish-yellow part of the spectrum. McKeon and Wright also state that the luminosity data vary little from observer to observer, and the mean curve is very similar to that of the protanope. In figure 7, due to Nelson, is shown hue discrimination of 5 deuteranomalous observers. Once again the general

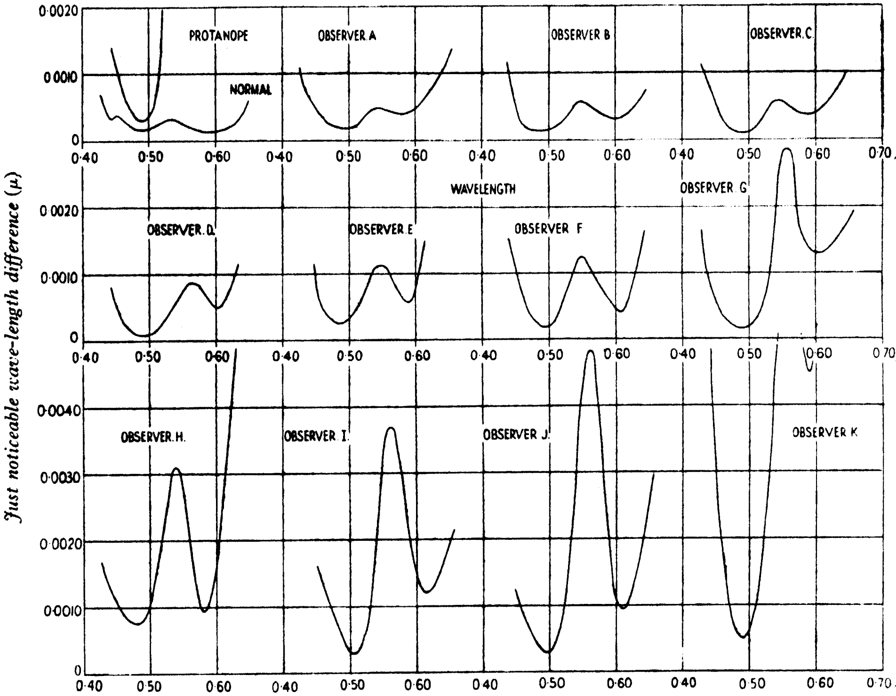


Figure 6. Hue-discrimination curves for 11 protanomalous observers arranged in order of increasing magnitude of defect. The first diagram shows the corresponding curves for the normal trichromat and for the protanopic type of dichromat (McKeon and Wright, 1940).

tendency is progressively poorer discrimination in the greenish-yellow. The luminosity curve of the deuteranomalous trichromats is similar to the deuteranopic curve.

Relation between dichromatic and trichromatic colours

The fact that dichromats can see all their colours as spectral colours has already been mentioned, and it is possible to show how the dichromat sees his colours in relation to the trichromat. In figure 8 is plotted the colour triangle for a normal observer, using red, green and blue primaries. The standard white is plotted in this diagram as a star, W. The dichromat sees white as a spectral colour, which is called the neutral point, and which is indicated in figures 2 and 5. For the average deuteranope this is at $\lambda = 0{\cdot}500\,\mu$. Therefore if the dichromat mixes a standard white and the colour at $\lambda = 0{\cdot}500\,\mu$, he will get a series of whites which will differ only in intensity. Therefore the straight line AW, in figure 8, which may be continued to point C, represents a locus of dichromatic white points. The point C gives us the equation

$$0{\cdot}500\mu = W = 0{\cdot}68R + 0{\cdot}32B, \qquad \ldots\ldots(1)$$

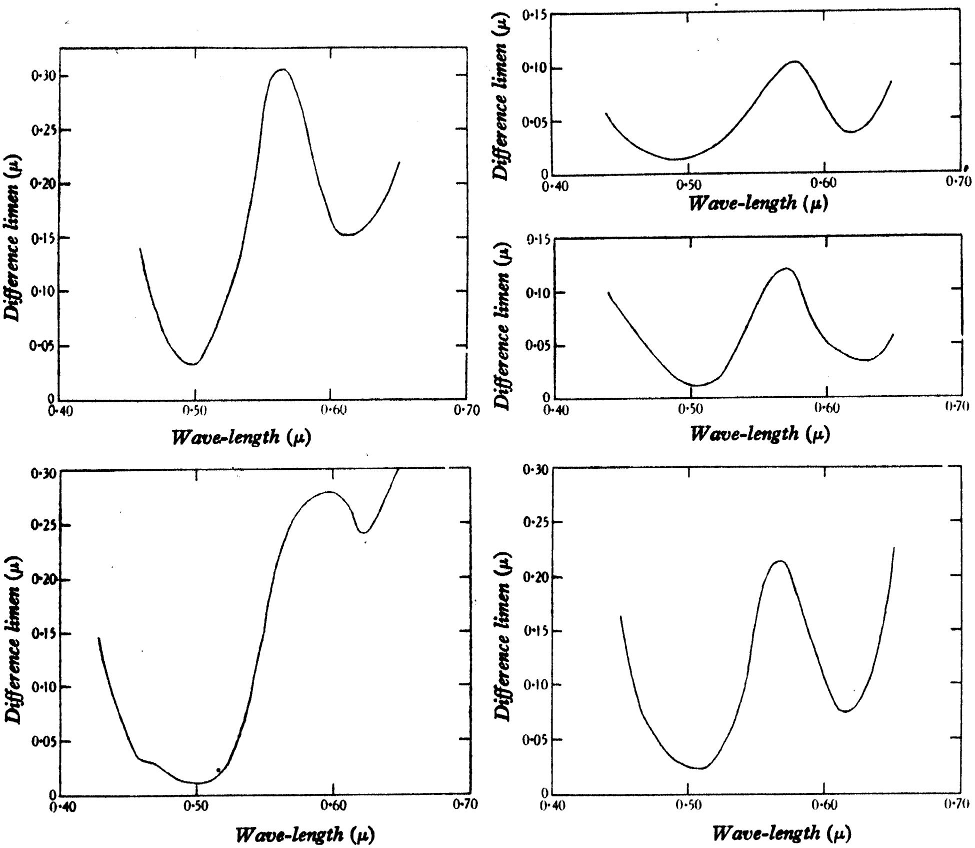


Figure 7. Hue-discrimination curves for 5 deuteranomalous observers (Nelson, 1938).

but from actual dichromatic matches we know that in terms of equal but arbitrary luminosity units

$$500\mu = W = 32{\cdot}26R + 1{\cdot}104B. \qquad \ldots\ldots(2)$$

Therefore, if we multiply the number of units of B in equation (2) by 13·75, we obtain the same relation between the R's and B's of equations (1) and (2), and thus establish a direct connection.

From the deuteranopic hue-discrimination curve we know that the discrimination at $\lambda = 0{\cdot}50\mu$ is $0{\cdot}002\mu$, and consequently, from $0{\cdot}498\mu$ to $0{\cdot}500\mu$

will represent one colour step. From the matching data we know how the deuteranope matches the colour at $\lambda = 0{\cdot}498\,\mu$ in equal, but arbitrary, luminosity units, and by multiplying the B value by 13·75 we know the value of this colour in terms of B and R on the green zero. This is point D. Starting from a convenient point, i.e. the white point, the boundaries of all the colour steps in the deuteranopic spectrum can be found, and the loci of these wave-lengths calculated and plotted in the colour triangle as shown in figure 8. The zones bounded by these lines contain, for the deuteranope, colours of the same hue. Figure 9 shows a similar diagram plotted for the protanope. Forty-three well-known paint colours, plotted in the diagram, indicate how the dichromats see these colours from the viewpoint of hue. It should be remembered, however, that the dichromat will only confuse colours in the same zone when the intensities, as seen by him, are of the same order.

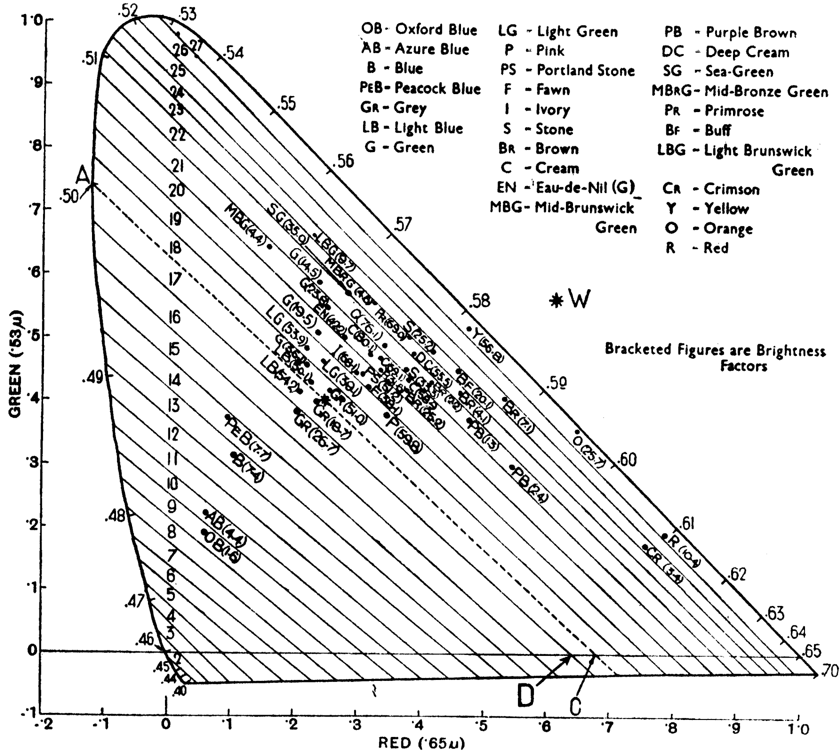


Figure 8. Iso-colour chart for deuteranopes (Pitt, 1935).

Relation between normal and anomalous sensations

The colour triangle for the dichromat reduces to a straight line, and as it is known, from actual measurement, that the number of colours seen by the anomalous trichromat is greater than the number seen by the dichromat, but less than the number seen by the normal trichromat, it is to be expected that the colour triangle, if plotted in exact relation to the normal colour triangle, should be smaller if the area is to represent the actual number of colours seen. If the colour triangles are plotted relative to the same physical stimuli it is obvious that these triangles will be of the same size, and it is therefore necessary to find some connection which will enable this triangle to be plotted in terms of the same subjective stimuli.

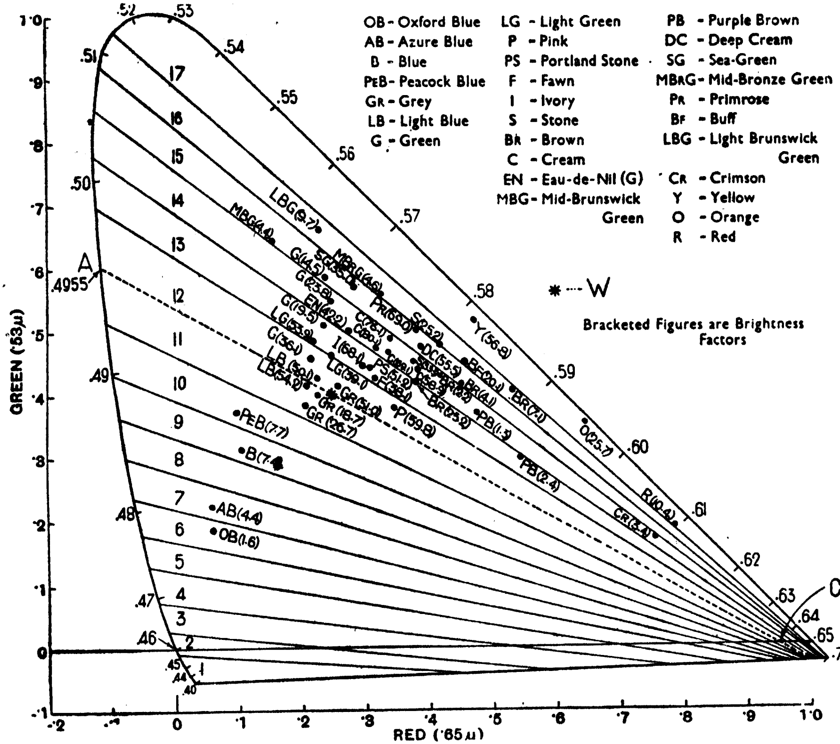


Figure 9. Iso-colour chart for protanopes (Pitt, 1935).

Nelson (1938) has discussed this possibility at length. He has assumed a previous suggestion put forward by Wright that the anomalous dichromat sees red and green as a desaturated orange and a yellow-green respectively. This is in agreement with hue-discrimination and saturation-discrimination data. From this it follows that the sensation of yellow is approximately the same for both the normal and anomalous trichromat. He further assumes, on the basis of McDougall's evolutionary theory, that the blue sensations are also the same in both cases. The points O and P, figure 10, represent blue and yellow respectively, and are seen by both the anomalous and normal observers, subjectively as the same blue and yellow. If it is supposed that one hue-discrimination step produces the same change in colour both for the normal and for the anomalous trichromat it is justifiable to assume that the number of steps for the anomalous trichromat between yellow and red, at $0{\cdot}65\,\mu$, corresponds in colour sensation to a colour distant from the yellow by the same number of hue-discrimination steps as for the normal observer. In this way it is possible to find a point on the normal colour-locus which corresponds to the colour seen by the anomalous observer at $\lambda = 0{\cdot}65\,\mu$. In a similar but less justifiable way it is also possible to find another point on the normal colour-locus which corresponds to the colour seen by the anomalous observer at $\lambda = 0{\cdot}53\,\mu$.

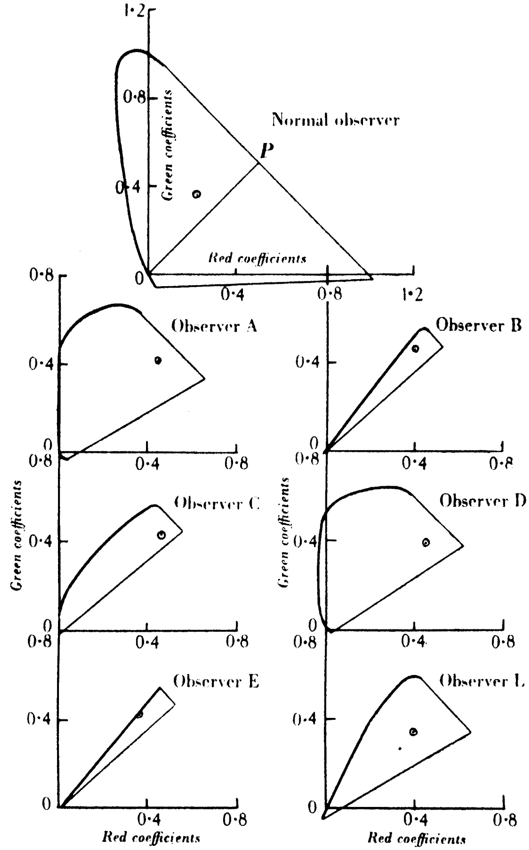


Figure 10. Reduced colour-triangles for 6 deuteranomalous observers (Nelson, 1938).

In each case the number of colour steps between yellow ($\lambda = 0{\cdot}5825\,\mu$) and red ($\lambda = 0{\cdot}65\,\mu$) and yellow and green ($\lambda = 0{\cdot}53\,\mu$) were calculated from the individual hue-discrimination curves and from the new coefficients for the red and green primaries, the coefficients of the blue primaries remaining constant; modified coefficients for the remaining wave-lengths were determined by ordinary transformation.

These modified coefficients are shown plotted in figure 10 and represent merely the colour triangle of the anomalous trichromat on a new set of primaries. Whether they represent, when viewed in relation with the colour triangle of the normal trichromat, a true subjective connection between the two systems is dependent on the validity of the assumptions made. Further evidence put forward by Nelson in some cases supports the results obtained, but in other cases disproves them.

The naming of colours

We do not know with any certainty what colours people with abnormal colour-vision see, but we do know that they see colours differently from the normal person. It is quite certain that the dichromat, with his limited range of colour and with the vast number of colour names at his disposal, makes use of different degrees of brightness to attempt to differentiate colours which to him are the same hue. To a more limited degree the anomalous trichromat adopts the same procedure. Consequently the naming of colours under haphazard conditions can lead to erroneous conclusions. In an actual test a dichromat named various decreasing intensities of a green ($\lambda = 0{\cdot}53\,\mu$) as yellow, yellow,

Table 1

Colour	Wave-length	Remarks by the dichromats				
		D_1	D_2	D_3	P_1	P_2
Red	0·6585–0·7800 μ	Yellow	Orange	Yellow-green	Dark colour —red ?	Orange
Orange	0·6000–0·6585	Yellow (warm)	Yellow	Orange-green	Yellow	Lemon
Golden-yellow	0·5828–0·6000	Yellow	Yellow	Orange-green	Yellow	Lemon
Yellow	0·5780–0·5828	Yellow	Yellow (orangy)	Orange-green	Yellow	Yellow
Greenish-yellow	0·5672–0·5780	Yellow	Yellow (orangy)	—	Yellow	Yellow
Green	0·5237–0·5672	Greenish-yellow	Yellow-white	Yellow	Yellow nearly gone, may be red ?	Light yellow to white
Blue-green	0·5020–0·5237	Whitish	Pink	Yellow	Yellow gone, nearly green	Blue
Blue	0·4306–0·5020	Blue	Mauve	Blue	Green	Royal blue
Violet	0·3900–0·4306	Blue	Blue	Blue	Dark colour	Violet

reddish-yellow, red, red. The preceding table shows how various spectral colours were named by three deuteranopes and two protanopes.

Another interesting demonstration of the fact that colour-naming is a bad test was given by a protanope. He was shown a coloured plate in a book, containing three distinct colours, a red, green and yellow, and he was asked what colour the red was. His first answer was green, but within a second he had changed his mind and called it red. After his first answer he had allowed his eyes to wander on to the other colours, and although they look the same colour to him he could discriminate between them by means of their intensity difference. The brightest he called yellow, the next brightest green, and the faintest red. If it were possible for a dichromat (say a deuteranope) to name colours correctly, he should say, on being confronted with the spectral colour at $\lambda = 0{\cdot}51\,\mu$: " Among other colours it is a spectral green, a mid-Brunswick green, an eau-de-nil green, a cream, a pink, a brown, and a reddish purple " (see figure 7).

Although the naming of colours by persons with abnormal colour-vision is not so bad as would be expected from the hue-discrimination data, this is chiefly due to the remarkable faculty that they have of using the brightness property to estimate colour. Indeed, some anomalous trichromats spend a large part of their life without knowing their colour-vision is abnormal.

Qualitative tests for colour blindness

Various tests are used which have been designed to detect colour blindness. All these tests, with the exception of the Nagel Anomaloscope, present to the observer confusion colours chiefly in the form of reds and greens (see figures 8 and 9) of such intensity as to render these colours indistinguishable to the dichromat and barely distinguishable to the anomalous trichromat.

The Wool Test, originally designed by Holmgren, was improved upon by Abney. This improved apparatus has 6 test skeins and 30 confusion skeins, which are so mounted as to present only one test at a time. Used properly, this test usually detects the dichromat but seldom detects the anomalous trichromat.

The Stilling Charts consist of a number of irregular and vari-coloured dots printed in such a manner that certain numbers may be seen by the normal observer but not by the dichromats or extremely anomalous trichromats.

The *Ishihara Test*, which is really an improved form of Stilling Test, is designed so that coloured dots, forming numbers in the diagrams, are easily distinguishable by the normal observer, but cannot be distinguished or are misread by colour-blind persons. Some of the tests can be seen as numbers by the dichromats, but are not so apparent to the trichromats. Other cards include both the above tests. This test, which should be used under daylight illumination, is easy to use and quick in operation. It picks out dichromats with unfailing regularity, but despite claims to the contrary, does not always distinguish between protanopes and deuteranopes. It picks out the more pronounced of the anomalous trichromats.

In the *Board of Trade Lantern Test*, small coloured apertures are viewed at a distance of 20 feet and the examinee is required to name the colours. All the lights exhibited are of equal brightness to the normal eye, and as the luminosity curves of both the trichromats and deuteranopes are approximately equal, the latter are readily detected when greens and reds are present. Neutral filters, to alter brightness, are introduced when a suspected protanope or protanomalous observer is being tested. The test, which should be performed in a darkened room, approximates to field conditions and is used extensively for testing of personnel in the Royal and Merchant Navies, the railway companies, etc. A lantern, similar to the Board of Trade Lantern, has been described in detail by Martin (1939). This lantern, when correctly used, provides much information on the actual quantitative aspect of colour deficiency.

The *Nagel Anomaloscope* is an optical instrument which will detect and differentiate between protanopes and deuteranopes and between protanomalous and deuteranomalous trichromats. The observer sees a small field of view, divided into two parts, one half being composed of a variable mixture of red and green and the other half a yellow of variable intensity. Dichromats can match a full red or a full green with a yellow by adjusting the intensity of the yellow. The amount of red used will readily indicate whether the dichromat is a protanope or deuteranope (see luminosity curves, figure 2).

The anomalous trichromat will match the yellow with either a reddish yellow or a greenish yellow, according to whether he is protanomalous or deuteranomalous, and the actual proportions of red and green will be an indication of the type and amount of abnormality. While the anomaloscope may be regarded as one of the best tests for colour vision, it suffers from the disadvantage that a certain amount of skill is required on the part of the observers.

Colour in industry

From what has already been said, it is obvious that in any trade or profession where colour discrimination or measurement is a part of his duty, the colour-blind person may be a perfect example of a square peg in a round hole. A product may be ruined by the employment of such a person, and in certain cases, which will be indicated later, lives may be endangered. Statistical measurements

Table 2. Results of statistical survey (males)

	Nelson	Schmidt	Waaler	v. Planta
Total number	1338	6863	9049	2000
Protanopes (per cent)	1·27	1·09	0·88	1·60
Deuteranopes (per cent)	1·20	1·97	1·03	1·50
Protanomalous (per cent)	1·27	0·68	1·04	0·60
Deuteranomalous (per cent)	5·08	4·01	5·06	4·25
Total deficients (per cent)	8·82	7·75	8·01	7·95

show that about 8 per cent of the male population are in some way colour-deficient. This is shown in detail in table 2, due to Nelson (1938).

Less than one in ten of the colour-deficient are female.

A general indication of the types of difficulties against which the person with abnormal colour-vision is always striving are given below:—

A car driver, although not bothered by the ordinary coloured traffic lights, could not distinguish between the green and red lanthorns used for controlling traffic when road repairing is in progress (see figures 8 and 9). He always had to wait for some other car to come along so that he had someone to guide him through. The reason why he could distinguish the ordinary traffic lights was because he knew which position the " stop " and " go " lights appear, and also because the so-called green light is really a blue-green. He would be further helped if, instead of the three lamps being circular, they were, say, square, triangular and circular.

Another case concerns the unfortunate position of a physics graduate when he was taking his intermediate examination in chemistry. He was performing a titration experiment in which methyl orange was used as an indicator. To the trichromat, methyl orange is orange in alkaline solution and turns red when the solution is made just acid. To the dichromat it does not change colour in any way (see figures 8 and 9). I remember the vivid description given me by the person concerned of the relief he felt on passing this examination and finally severing his connection with methyl orange. Another example is more amusing than distressing. The person concerned stated that he could distinguish green cherries on a tree but he could not see red ones. He was a protanope, and the red cherries appeared almost black to him, and in the dark shadows of the leaves were almost invisible (see figure 3). The instance of a mourner appearing at a funeral wearing a red tie is well known. A copy of figure 3 would absolve him from all blame.

It is difficult to make a full list of all trades and professions and state specifically where colour blindness would prove detrimental. It is possible, however, to make reference to some of them, which may serve as an indication for all. Personnel for the Royal and Merchant Navies are subjected to very stringent tests. This is to enable those responsible to ensure that men who are colour-deficient are not placed in positions which would render them potential sources of danger. The look-out man must be able to differentiate navigation lights up to two miles in clear weather, or at safe distances in haze, rain, fog or snow. He must be able to distinguish between these lights solely by their colour quality. At sea, the dichromat as a look-out man would be dangerous, but it should be remembered that the anomalous trichromat, who can often differentiate between the colours when given time to deliberate, would cause delays which again might be dangerous. It is stated (Martin, 1933) that the anomalous trichromats present greater danger, due to the fact that they are more difficult to detect. For similar reasons, personnel for the railway companies are also rigidly tested. Those who are colour-deficient are not employed on duties which necessitate the reading of coloured signals.

Soldiers employed in the Signal Corps and certain members of the R.A.F. are obvious cases where only normal trichromats should be used. The latter make use of coloured lights in one way or another, but it should also be realized that a pilot, when making a forced landing, has to make use of his ability to discriminate colours in order to ascertain that the particular spot on which he chooses to land is the safest one in the vicinity.

The four cases just described are ones where persons who cannot readily distinguish colour are a source of danger. In other industries or trades, persons may be required not only to distinguish colour but also to create and directly match them. These persons may not endanger life, but they would certainly prove to be a liability. The paint and varnish industry is a good example of such a trade. Mid-Brunswick green and some types of brown paint are of almost equal brightness to the dichromat. He sees them as the same colour (figures 8 and 9) and would match them without hesitation. It is not beyond the bounds of possibility for him to match an actual crimson paint with a light Brunswick green. Such glaring examples are merely instances of the capabilities of the dichromat and extreme anomalous dichromats, and would probably be readily detected. It is when paints of smaller differences are matched (and here the majority of the anomalous trichromats enter the field) that the colour deficient prove themselves a real nuisance. It goes without saying that such people employed as colorimetrists are worse than useless. From the colour viewpoint, that which applies to the paint industry similarly applies to other such industries as woollens, textiles, dyes and inks, and also to persons retailing these products.

A special mention should be made of the paper and other industries where the measurement of whites are of utmost importance. It has been stated that measurements of the colour coefficients of whites should be made to an accuracy of at least one in the third decimal place, to satisfy the requirements of the users. The dichromat would distinguish himself by getting the first decimal place wrong and even then not being consistent, while the anomalous dichromat would probably get the second decimal place wrong. Even persons with a slight colour deficiency would be incapable of making a correct normal reading. It has been generally realized by those directing industries where extreme accuracy of colour matching or measurement is required, that only observers possessing closely normal colour-vision should be employed in certain positions. Excessive variation of the macular pigment from the normal would also introduce errors too large to be acceptable, and tests which measure the Y.B. ratio have been devised with the view to detecting persons who, although possessing normal trichromat vision, suffer from this defect. This test, suggested by Ives and Kingsbury (1915), consists in comparing the amount of light transmitted by two filters, one yellow and the other blue. The ratio of the brightnesses for the normal eye can be calculated from a knowledge of the colour temperature of the source and the spectro-photometric curves of the filters. The extent of the departure from this ratio indicates the degree of abnormality of the observer. In fairness to all these people, however, it should be placed on record that spectro-photometric readings, being purely physical measurements, could be correctly made by all classes of colour-blind observers. Even so, it should be noted that, due to the low luminosity in the red, the protanopes and protanomalous trichromats would not be able to measure the intensity of the longer wave-lengths with any degree of accuracy.

The manufacture of cosmetics has recently developed into something approaching a major industry. Women, with their very keen sense of discriminating even the most subtle difference in colour, would be very quick to notice a coloured preparation created by a colour-deficient person.

The ability to discriminate small differences in colour is often of great value to the practical physicist. He would find it quite impossible to determine the colour temperature of a lamp visually with the required degree of accuracy if his colour vision was in any way abnormal, because his discrimination of colours would not be good enough. Manufacturers of colour-temperature meters, the use of which depends on the blending of a green and yellow to match the colour of the subject, state specifically that the apparatus should not be used by colour-blind observers. As users of a flicker photometer, for the measurement of heterochromatic colours, their accuracy as compared with the standard observer would depend upon whether they were protanopic or deuteranopic. If of the first class, whether a dichromat or anomalous dichromat, an observer would give erroneous readings, the magnitude error of which would depend on the nature of colours being matched. This error would, in the case of the protanomalous worker, also depend on the extent of his colour deficiency. The error may be calculated from the luminosity curve of the actual observer and the spectrophotometric curves of the colours to be matched (if known). Due to the luminosity curves of the deuteranopic and deuteranomalous observers being similar to the standard curve, the differences in these cases are not so marked. The extent to which colour is important to the spectroscopist is obvious. The method of measurement adopted and the great use made of photography protect him to a large extent if he is colour-deficient. Colour-staining of microscopic sections can be misleading to the colour-deficient microscopist if the colours produced are red, yellow, or green.

The use of methyl orange as an indicator has already been mentioned. The chemist also makes use of flame tests which indicate to him the nature of the substance under test, and many chemical reactions are characterized by a change in colour, often a small one. While it would be foolish to assume that a colour-blind observer could not follow the profession of chemistry or physics, it is certain that his usefulness is restricted in certain branches of these subjects, and awareness of the nature and extent of the defect would be of use to him.

What has been said about the chemist also applies, although to a lesser extent, to the pharmacist. Members of the medical profession also rely to a degree, which is probably greater than they realize, on colour as an aid to diagnosis.

The photographic industry provides another outlook on the subject. Many of the sensitized goods have to be processed in red light. Protanopes and the more extreme cases of protanomalous trichromats would find it almost impossible to work under such conditions. It should be realized, however, that colour-blind persons would have no difficulty in working under panchromatic (green) safe-lighting conditions, as none of these persons, not even the so-called green-blind, have luminosity deficiency in the green. The obvious conclusion from the above remarks is that only protanopes and protanomalous trichromats need be debarred from working in dimly illuminated darkrooms, and only then when red lighting is employed. The preparation of colour films is another concern of this industry. As it is impossible, for reasons too involved to be discussed in this paper, to reproduce actual colours accurately, it is necessary to reproduce them in such a way that the complete picture, as a whole, looks correct. Although the individual colours may depart markedly from the original colours, and the whole still be accepted as a good reproduction, what may be termed the balance of colour must be fairly exact to obtain the effect. Many tests and much checking against standards must be effected by competent personnel before high-class colour pictures are obtained, and it is evident that the colour-vision of this personnel must be normal in every respect.

What has already been said about colour-films also applies to the work of the interior decoration artist. This artist must be able to present a good " colour scheme ", and while it is not everyone who could follow such an occupation with success, it is obvious that the colour-deficient person would never be able to perform such duties successfully except for persons of the same degree of colour deficiency as himself. It is here that colour is really treated subjectively, being regarded in contrast with other colours, and many of the industries described must employ persons capable of treating colour in this manner. For instance, the person creating a primrose-yellow paint must produce a paint the colour of which resembles the colour of the primrose when seen in contrast with its dark green leaves. In fact he produces, objectively, a more saturated yellow which, when seen on a wall and without the contrasting green, looks like the yellow of the primrose.

It is not only in the trades directly concerned in the manufacture of coloured goods that persons with normal colour-vision should be employed (or rather, from which persons with abnormal vision should be excluded). Most people have passed pleasant times watching workmen sitting in a small hole in the pavement quickly and nonchalantly joining thousands of telephone wires together, and have marvelled how they managed it. The telephone cables, sometimes containing as many as two thousand pairs of cables, are colour-coded. The pilot wire, from which the start is made, is usually coloured orange, and spiralling from the pilot are the other wires, which are coloured somewhat differently in the different-sized cables. In one of the smaller types of cable, the colours are white, black, blue, green, red, red and black, blue and black, green and black. It is known that colour-deficient persons have difficulty in joining these cables, but providing they keep the order of wires correctly they can accomplish the jointing. With the larger cables, which contain a more complicated colour-system, they find it correspondingly more difficult, and often, when trying to locate a single pair of wires on which a fault has occurred, they find the task beyond their capabilities and are forced to seek advice from the mate who is invariably helping in the operation. The obvious moral is that if the jointer is colour blind, his mate must have normal colour-vision.

Many factories have colour-coded supply systems. Thus A.C. electricity, D.C. electricity, hot and cold water, gas, steam, brine and air are housed in pipes marked in some way with a distinguishing colour. Many of these pipes are similar in size, and while it is not suggested that the colour-deficient person would unwittingly connect, say, the gas and water together, it might occur that, after cutting off the steam system in order to repair it, he would then proceed to take the water pipes apart to the disadvantage of all concerned. What is probably more important is that other minor troubles, not always brought to the notice of the authorities, are almost certain to occur which cause delay of some description.

It is hoped that, from what has been said on the subject, the need for testing the colour vision of all persons connected with certain trades or professions will be realized. It is, however, obviously to be desired that the tests should not be made after the person has commenced employment, as his subsequent removal might lead to considerable economic hardship. The tests might be made before entering a given occupation or, better still, before the future employees have left school. This would then allow time for a more appropriate choice of employment to be made; it might prevent training being given that subsequently turned out to be useless; it might save the child a certain amount of ridicule, embarrassment and difficulty while still at school. The tests would, presumably, have to be made by the medical officer, but it should not be overlooked that the science master or the art master would almost certainly be interested in the test, and his co-operation might help considerably to reduce the labour. In general, the tests of the Ishihara type are best where mass testing is required. Where the defect is likely to cause danger, a further test must be adopted to ensure that all the colour-deficient persons are weeded out; for this, the Board of Trade lantern test is usually used. Personssuspected of being colour deficient should, if possible, be further tested to ascertain whether they are dichromats or anomalous trichromats, and whether they are of the deuteranopic or protanopic types, for often (as in the case of deuteranopes and deuteranomalous trichromats in red-illuminated darkrooms) some classes of colour-deficient persons can work with the same efficiency as normal trichromats even when colour plays an important part in the scheme of things. This dividing of colour-deficient persons into their various groups can only be performed with any degree of certainty by using an apparatus of the Nagel-anomaloscope type. As already stated, this apparatus in its present form is difficult for an unskilled observer to use, and its re-design to obviate this rather formidable difficulty should be one of the aims of those responsible for the testing of colour blindness. It is further hoped that one day the various responsible authorities will cease to colour-code in the so-called confusion colours, as in the green and red lanthorns of the road repairer, and will do more to mitigate the lot of the colour deficient, in the same way as a golfer acquaintance, a protanope, who realized that, because of his low luminosity in the red, he could not see red golf-pegs, and consequently lost them, solved his difficulty by using yellow ones.

REFERENCES

ABNEY, W. DE W., 1913. *Researches in Colour Vision and Trichromatic Theory* (London).
GIBSON, K. S. and TYNDALL, E. P. T., 1923. *Sci. Pap. Bur. Stand., Wash.*, no. 475.
GUILD, J., 1924–5. *Trans. Opt. Soc.* **33,** 73.
GUILD, J., 1931. *Philos. Trans.* A, **230,** 49.
HECHT, S. and SCHLAER, S., 1936. *Amer. J. Physiol.* **20,** 57.
IVES, H. E. and KINGSBURY, E. F., 1915. *Trans. Illum. Engng Soc., N.Y.*, **10,** 203.
JONES, L. A., 1917. *J. Opt. Soc. Amer.* **1,** 63.
KONIG, A., 1903. *Gesammelte Abhandlungen zur physiologischen Optik* (Leipzig).
KONIG, A. and DIETERICI, C., 1892. *Z. Psychol. Physiol. Sinnesorg.* **4,** 241.
LAURENS, H. and HAMILTON, W. F., 1923. *Amer. J. Physiol.* **65,** 547.
MARTIN, L. C., 1933. *Spec. Rep. Ser.* (Med. Res. Counc., Lond.), no. 185.
MARTIN, L. C., 1939. *Brit. J. Ophthal.*
MCKEON, W. M. and WRIGHT, W. D., 1940. *Proc. Phys. Soc.* **52,** 464.
NELSON, J. H., 1938. *Proc. Phys. Soc.* **50,** 661.
PITT, F. H. G., 1935. *Spec. Rep. Ser.* (Med. Res. Coun., Lond.), no. 200.
STEINDER, O., 1936. *Sitzungsber. Akad. Wiss. Wien,* **115,** 39.
WRIGHT, W. D., 1927-8. *Trans. Opt. Soc.* **29,** 225.
WRIGHT, W. D., 1929. *Spec. Rep. Ser.* (Med. Res. Coun., Lond.), no. 139.
WRIGHT, W. D. and PITT, F. H. G., 1934. *Proc. Phys. Soc.* **46,** 459.
WRIGHT, W. D. and PITT, F. H. G., 1935. *Proc. Phys. Soc.* **47,** 205.

DISCUSSION

Dr. SHAXBY. I am glad that Dr. Pitt treated the subject from what he called the physical standpoint, i.e. of the experimental investigation of the facts of colour vision and its abnormalities; theories of colour vision are not yet such as to help us in the practical questions of the disabilities for various professions that we are considering to-day. We still require more statistics, and particularly a good quick method of assessing trichromatic anomalies. Colonel Ishihara's test still seems the most satisfactory preliminary one for general use.

I think the Colour Group can serve a very useful purpose in industry: we know all too little of the occupations in which colour deficiencies are a bar, or at least a handicap, by slowing up the work.

To the professions Dr. Pitt listed we may add that of the geologist, with his special maps, coloured to indicate the various strata.

I consider that it is very desirable that tests of colour vision should be made in schools, early enough to prevent boys preparing for occupations for which colour blindness will unfit them, as soon, in fact, as they are of age to make the tests satisfactorily. In secondary schools I think the test should be made at the time of the initial medical examination; in primary schools not later than at the last medical examination. In one school alone, in which I am interested, in the last few terms two boys have been rejected for the R.A.F. and one for the Merchant Service; not one of them had any suspicion that his colour vision was not normal.

I hope that a special sub-committee of the Colour Group may be formed, whose duties shall include, among other things, the obtaining (from the various industries interested in questions of colour) of information, as complete as possible, of the techniques and processes in which colour deficiency is a bar, and particularly those in which it causes slowness or uncertainty.

Dr. C. **Maunder Foster**. The accuracy of colour vision is to some extent a function of training, and the degree of colour blindness might be more accurately measured than at present; standard limits might be established between which colour perception might be considered adequate for various purposes. The marked discrepancy in relative incidence of colour blindness between male and female may indicate some connection between the gonadotropic hormones and the mechanism of colour vision.

The accurate use of indicators in analytical work is an instance of an important case where slight errors in colour vision gave markedly different results. I should like to ask whether slight errors in colour vision could in some measure be compensated by altering either the colour temperature of the illumination or by filters.

Dr. **Strange**. I wish to support the suggestion that there should be a regular colour-vision test of all children. The need for this test is usually underestimated because of our feeling that any anomalies in vision can hardly pass unnoticed, at least among scientific workers. This view ignores how essentially normal even a dichromat's colour perception is to himself, and it is only unusual circumstances, such as green berries on a holly tree, or special tests, such as indicator end-points, that disclose any abnormality. This may be illustrated by my experience among one group of about 50 scientific workers, where three hitherto unestablished, if not unsuspected, colour-blind persons were discovered. The first two were anomalous trichromats, one a physicist of over 30 and the other a laboratory assistant; the third was almost a dichromat, and had been given the job of testing cathode-ray tubes, including the colour of the screens !

These examples indicate another essential feature of the proposed tests, It is not enough to sort children out into those having standard and anomalous vision; subsidiary tests must be established to define the degree of anomaly and limits set up to indicate what degrees are likely to be a serious handicap in different occupations. In two of the cases mentioned, the anomalous vision caused only slight inconvenience, and was in no sense a bar to success in the chosen occupations.

Dr. **Walsh**. Is the Ishihara test sufficient as an indication of suitability or unsuitability for such occupations as the Navy or the railway service, or would it not be necessary to use a lantern test as well?

Mr. **Preston**. I would like first to ask a question, to make certain that I have properly understood part of Dr. Pitt's very interesting discourse. Is it correct that, among defective observers, the poorer are the powers of hue discrimination, the farther the luminosity curve is displaced from the normal or average? In other words, is it possible for some observers to have a luminosity curve approximating to the average or " standard " eye and yet be deficient as regards recognition of colours?

I suppose it is difficult or impossible at the present stage to indicate what are the absolute, as distinct from relative, sensations produced by light of different wave-lengths for defective eyes. The curves of luminosity function on a *relative* basis show the differences in sensitivity to radiation of different wave-lengths for the same eye, but do not necessarily provide data for comparing the brightness sensations produced in two *different* eyes by the *same* monochromatic source.

Dr. Pitt has hinted that the low values of the relative luminosity function for the protanope in the red indicate a low *absolute* sensitivity, since the protanope is unable to see at all by the light of a red photographic dark-room lamp. This suggests that curves of threshold values of illumination, or, rather, radiation, throughout the spectrum would be interesting; but perhaps I am anticipating too much in this remark.

Dr. W. D. **Wright**. I should like to support Dr. Pitt in his belief that in many industries colour blindness must be a handicap both to the individual and to the firm concerned, and I think there can be no doubt that testing for colour blindness in such industries is very much to be desired. We have to remember, however, that those people who were found to have defective colour-vision might be exposed to considerable economic hardship if they were shown to be unfitted for their particular job; and any campaign to encourage testing in industry would have to be accompanied by assurances that there are a great many activities where colour blindness is no bar to efficiency and to which defective observers could be transferred. From this point of view, the suggestion that tests should be made in schools is excellent, since it would sort out the colour-defective cases before they entered industry and would avoid any suggestion of subsequent hardship.

There are a number of practical questions that arise in connection with testing at schools, and I have put certain of these to Dr. Gale, of the Board of Education, as he was unable to be present when Dr. Pitt read his paper. They are dealt with in the following contribution, but Dr. Gale has asked me to say that any opinions expressed are his own and do not necessarily represent those of his Department.

On one point, Dr. Pitt's paper is slightly misleading because the information is incomplete. Physicists generally simplify colour problems by reducing them to two variables instead of three, by regarding brightness differences as being distinct from colour differences. Thus, figures 8, 9 and 10 are really plane sections through the colour solid, and when Dr. Pitt says that a trichromat can distinguish about 5000 colours with a 2° field, this refers only to the plane diagram and not to the colour solid. From the point of view of the dyer, the painter and the man in the street, a brown is a different colour from a yellow or orange, although they may be represented by the same point in the colour triangle, and when brightness differences are included, the number of distinct colours runs into millions. To get a complete picture of the colour perception of the colour-blind observer, we ought, therefore, to consider what happens to the trichromatic colour solid (as described, for instance, by Newhall (1939)*) in the case of the anomalous trichromat and dichromat. Presumably, in the case of the anomalous observer, the solid is reduced in volume in various directions according to the type of defect, while the dichromatic solid becomes a plane. Has Dr. Pitt any information on this question?

Dr. GALE. I should like to submit the following notes in reply to questions which Dr. Wright has sent to me:—

1. In 1938, 1518 doctors were employed on school medical work. The majority of these did other work, usually in some other branch of the Public Health Service, as well, and the total amount of time which they gave to school medical work was equivalent to that of about 728 whole-time doctors. This was about 1 to 6220 children in average attendance.

2. Medical inspections are divided into (*a*) routine, (*b*) special—when the child is brought forward for examination by parent, teacher or nurse, for some special reason, (*c*) re-examination of children found at a previous inspection to need special care. The total numbers of inspections in 1938 were:—

Routines	1,677,008 (37 % of total number in average attendance)
Specials	1,563,917
Re-examinations . . .	2,182,157

---

* S. M. Newhall, 1939. *Amer. J. Psychol.* **52**, 383.

3. In elementary schools, routine inspections are usually carried out on entrance (about age 5), at about age 8, and at about age 12 (some areas have four inspections). In secondary schools, procedure varies considerably, but generally four full routine inspections are done during a child's school career. The conditions of working vary from a screened-off corner in a classroom of a village school (with lamps as the only artificial light) to a well-appointed inspection room in a modern school. Usually about 5 to 8 minutes are allotted to each child. This time varies very much from child to child, but if more detailed examination is necessary the child is usually referred to a clinic. Vision is often tested beforehand by the school nurse, using Snellen's types, and she mentions to the doctor any child who appears to have a defect. Vision is usually tested only at the two later inspections because of the difficulty and unreliability of testing young children who do not know their letters.

Now for the answers to particular questions.

(1) *Would the introduction of a colour-vision test be feasible?* I think the answer to this depends very largely on what the test is. It would have to be quick, not too complicated, and capable of being carried out in an ordinary room, possibly without electric light. This is where the Colour Group could be most helpful, i.e. in the choice of possible tests.

(2) *Would it be better to test elementary or secondary school children, or both?* If a test is thought desirable, I think it should be carried out for (*a*) secondary-school children at their last examination, and (*b*) elementary-school children, also at their last examination. The answer to this question does, however, depend on a good many things, and I should not like you to regard it as anything but tentative.

(3) *Would it be the responsibility of the medical officer or the teacher to advise on suitable employment?* I do not think that the medical officer could know enough about industry to decide how far a defect would interfere. It is part of the medical officer's duty to give general advice on employment at the time of the last medical inspection. His advice is usually with reference to things like unfitness for heavy physical work or unfitness for work in a dusty atmosphere, etc. Such advice is, of course, given in consultation with the head teacher.

(4) *Would the Board approve of testing by the science master?* I do not think I can answer this, as it raises some rather thorny questions. From the purely medical point of view, I see no objection to the test being done by anybody who is competent. It must be remembered that there is no special science master in very many elementary schools.

(5) *Vocational guidance.* Neither the school medical officer nor the teacher could be expected to know the exact requirements of a given industry. I should have thought that the certifying factory surgeon was the man to deal with

this aspect. Another question which arises is—in how many jobs is a defect of colour-vision important, i.e., would it be justifiable to test the colour-vision of something over half a million children a year if only, say, 1 % or 10 %, or whatever the percentage is, were going to jobs where colour-vision mattered? Some estimate of the percentage of children entering such industries would obviously be important, and I confess I have not the remotest idea of what it really is. Here, again, I am sure members of the Colour Group would have useful information.

(6) *A.T.C. admission.* This is really a matter for the Air Ministry, and I am speaking rather without the book, but I gather that a defect of colour-vision is no bar to admission, because there are many jobs in the Royal Air Force where the defect does not matter. When candidates are transferred from the A.T.C. to the R.A.F. they go through the full R.A.F. medical examination, which, I believe, includes tests of colour-vision.

(7) If representations were made to the Board of Education I am sure they would be carefully considered, but I think an overwhelmingly strong case would have to be made out before anything new could be added to the work of school doctors, or teachers, in present circumstances.

Prof. L. C. MARTIN. I approve the suggestion that tests of colour vision in schools should be instituted when opportunity offers, but it will not be a simple matter to arrange the scheme. For example, it will not be sufficient to tell a lad: "You are an anomalous trichromat, and therefore cannot hope to undertake watch-keeping duties in the Navy." There must be an opportunity for a supplementary examination if required—and this supplementary examination would possibly have to be of a more or less official character.

But even under war conditions, some colour-vision testing might well be undertaken in secondary schools as a voluntary measure, and would tend to prevent disappointment while helping to make sure that effort will not be misdirected in preparation for forms of national service for which an individual is not suitable.

AUTHOR's reply. There is little to be said in reply to Dr. Shaxby beyond supporting him in his plea for the more intensive testing of school children and to say that the sub-committee to which he refers is already taking shape.

In reply to Dr. C. A. Maunder Foster, slight errors in colour vision can be compensated by the use of suitable filters. This is described in detail in the paper by Ives and Kingsbury (1915), to which I have already referred.

Dr. Strange's discussion emphasizes the need, which in my opinion is real, for adequate as opposed to superficial colour-testing.

In reply to Dr. Walsh, the Ishihara test is not considered sufficient as an indication of the suitability for such occupations as the Navy or the railway service. A lantern test is always used in addition to the Ishihara test.

In reply to Mr. Preston, there is no evidence whatever for assuming that poorer powers of hue discrimination are in any way connected with the displacement of the luminosity-curve. The deuteranopic luminosity-curve is very similar to the normal curve, and yet the discriminatory power of the deuteranope is much lower than that of the normal observer. Figure 6 shows an increasing magnitude of defect in the discriminatory powers of 11 protanomalous observers, but these observers have almost identical luminosity-characteristics. The same phenomenon is also shown by deuteranomalous observers. It is quite impossible to measure the absolute sensation curves of the eye, but by making various assumptions it is possible to indicate the possible shape of such curves. There is no doubt that the protanope has a low absolute sensitivity in the red, and that threshold measurements would supply interesting and useful data. To what extent these data would be used, in comparison with similar data completed for normal observers, is difficult to foresee, as once again " various assumptions " would have to be made.

I find Prof. Martin's query, as to what is to be done for a young person after he is told that he is colour deficient, difficult to answer. All I can suggest is that he should be told by a competent person what particular jobs should be avoided. As pointed out, it does not always follow that he should avoid occupations where use is made of colour.

The first part of Dr. Wright's discussion may be regarded as a partial answer to Prof. Martin's query. With reference to the colour solids, for a proper understanding of what I consider to be a difficult aspect of the subject, the paper, which took longer than the time limit set to deliver, would have tended to become interminable. I agree, however, that only the colour solid presents the true picture, but beyond agreeing with the assumptions which Dr. Wright makes, I have nothing to add beyond expressing the hope that one day somebody will perform the useful but arduous task of expressing the characteristics of the anomalous observers in this form.

Dr. Gale's reply to the questionnaire put to him by Dr. Wright forms a very useful basis on which to start on the problem of testing school children. Dealing with a query raised by Dr. Gale, there is no doubt that the Ishihara test fulfils his requirement of a speedy test capable of being used without complicated accessory equipment. A person considered competent to teach should be able to supervise the use of these tests. The question of vocational guidance is again raised, and it would seem that the solution of this problem will be an important task for the aforementioned sub-committee.

CANAL-RAY POSITIVE-ION SOURCES

By J. D. CRAGGS, M.Sc., A.Inst.P.,
Metropolitan-Vickers Electrical Co., Ltd.

Received 8 *January* 1942

ABSTRACT. The paper deals with the positive-ion output available from cold-cathode glow discharge tubes with a canal drilled through the cathode.

The importance of obtaining large canal-ray currents for high-voltage positive-ion bombardment work is stressed, and a brief preliminary discussion of previous work given. The reasons for adopting canal-ray sources are then discussed. A study of the discharge mechanism of such sources reveals several interesting features.

In the measurement of intense positive-ion currents, precautions were taken to avoid errors due to secondary emission of electrons from the collector, and, a satisfactory technique having been developed, exhaustive measurements were taken on canal-ray sources of two types differing in detail from each other.

The necessity for using a low gas output is stressed, and the effect on yield of varying the canal-flow impedance is studied. Some interesting results were obtained by the use of a focussing diaphragm in the source, e.g. yields were considerably increased, but with a higher tube pressure. Total ion currents up to 2 ma. in strength were obtained.

The presence in the canal-ray beam of large numbers of neutral particles was shown by an experiment on electrostatic deflection of the rays after leaving the canal.

The final part of the paper deals with magnetic deflection of the beam and its analysis into different groups of ions. This work was done with a high-voltage cascade acceleration tube, and qualitative data on collision processes in molecular hydrogen are given.

§ 1. INTRODUCTION

The production of focussed positive-ion beams of high current-density ($\sim$100 microamps/cm^2) is now a problem of extreme importance in the technique of artificial nuclear disintegrations. The inefficiency of such reactions must be offset as largely as possible by the use of large numbers of bombarding particles. For instance, 50 μa. of protons at 550 kv. cause an emission, from lithium targets, of gamma rays equivalent approximately to 0·5 mg. Ra (a yield of the order $7/10^{10}$ photons/proton). If such gamma rays are now used to cause neutron emission from a secondary target, there is a further large diminution of effective output, and the results are positive in these conditions for only a few substances (Bothe and Gentner, 1939; Chang, Goldhaber and Sagane, 1937). It is therefore desirable to produce focussed ion beams of 0·1 to 1 ma. at medium voltages ($\sim$500 kv.). As nearly all ion sources produce strongly divergent beams, the focussing problem is largely an independent one, and will be considered separately.

It is not difficult to produce a source yielding 1 to 100 μa. of positive ions, and almost any gaseous discharge device with a perforated cathode will give emergent currents of this order. The problem of increasing the output is not entirely one of increasing the power input to the source. Some of the difficulties are discussed below, with reference particularly to the type described by Oliphant and Rutherford (1933), i.e. the low-pressure, high-voltage, glow discharge-tube with a cathodic canal.

§ 2. TYPES OF ION SOURCES

With the exception of a few special devices, such as Franzini's source (1938), in which a palladium tube with an internal pressure of 1 atmosphere of hydrogen surrounded by a negatively charged electrode gave 100 to 300 μa. of positive ions (20 kv.) at about 10^{-5} mm. Hg pressure, positive-ion sources fall into two main classes: (*a*) low-voltage arcs, and (*b*) glow discharges. Before proceeding to the description of work carried out on sources of type (*b*) it is useful to draw comparisons.

Type (*a*), which is widely used in America, consists usually of an arc passing between a hot cathode (through a constriction) and a massive anode. The ions are drawn out through a hole leading radially into the constriction by means of a probe to which a relatively high potential ($\sim$5–10 kv.) is applied. In some tubes no probe is used, and ions and electrons diffuse out of a radial hole. The arc voltage may be 100 to 200 at a current of 1 to 2 amp., and a gas pressure (hydrogen or deuterium) of 0·01 mm. Hg. Some excellent results have been obtained with sources of this kind. Lamar, Buechner and Compton (1937), Timoshenko (1938) and Zinn (1937) obtained outputs of several hundred μa. at arc currents of 1 to 4 amp. ($\sim$100 v.). The capillary, through which the arc passed, was 3·5 mm. diameter.

In the case of the cold-cathode glow discharge sources, the experiments of Oliphant and Rutherford (1933) have formed the basis of all subsequent work. (See, for example, the experiments of Bouwers, Heyn and Kuntke, 1937.) Their source consisted essentially of two heavy concentric steel cylinders for the anode (inner) and cathode (outer), the latter being closed at one end with the cathode proper, which had a central hole bored in it. The cylinders, of length about 1 metre and diameter 12 cm., had a small sideways clearance (few mm.) to prevent stray discharges, the main discharge passing from the region inside the anode cylinder to the cathode block, from the hole in which the canal rays emerged. For 20 ma. discharge current at about 20 kv., a beam strength up to 1 ma. was obtained. It was stated that, at first, the beam consisted mainly of molecular ions but, after running, these broke up and the beam became composed principally of protons (hydrogen gas was used) and neutral particles. Sources of this type have become general in European practice. Bothe and Gentner (1937) and Amaldi *et al.* (1939) have used them, and obtained target currents of about 50 μa. in high-voltage acceleration tubes, i.e. their source output currents were probably about 100 μa.

Roberts (1937) has published a direct comparison between arc and canal-ray sources, also using a complete disintegration set. The canal-ray source gave 100 to 200 μa. output, but only 2 to 3 μa. of ions of mass one were focussed on the target. A high proportion, about 50%, of atomic ions was obtained, but it was considered that the power and voltage required were disadvantages. Oliphant (1935), comparing the two types of source, considered properly designed canal-ray tubes to be most suitable, and stated that arcs require, in general, hot filaments of short life, and pure gases. It also seems clear that arc sources are more complicated to build and operate than the glow-discharge tubes.

Finally, the comprehensive work of Hailer (1938) on canal-ray sources should be mentioned (see below for further references). The curves (loc. cit.) indicate an output current of about 1·5 ma. at 30 kv., 25 ma. discharge and a pressure of about 0·18 mm. Hg of hydrogen. The proton content was about 30%.

The above discussion ignores the problem of focussing the beam after it has left the canal. Both types of source give divergent beams, but as the glow-discharge sources in general run at a much higher voltage (20 to 50 kv.), correspondingly higher focussing voltages are required. This is not usually a serious defect, as such sources are almost invariably used with high-voltage (100 to 1000 kv.) apparatus, and potential division, either by corona points or resistances, is usually feasible.

§ 3. MECHANISM OF GLOW DISCHARGES AT LOW PRESSURES

In order to obtain as high a yield as possible from canal-ray sources it is necessary to consider the physical properties of the discharge. For the highest efficiency, the latter should be run near the low-pressure limit of such discharges, probably because the radial components of the high field-strength thereby encountered are effective in partially concentrating the positive-ion current before the cathode is reached. A larger proportion of the current is thus available for the formation of canal rays. It is found that the appearance of the discharge in these conditions is different from the normal régime, in that the various layers (cathode glow, cathode dark space, negative glow, negative dark space and positive column) largely disappear. Most of the tube becomes filled with the cathode dark space (Hailer, 1938), and the cathode glow extends outwards from the cathode and recedes from the tube walls, thus forming a thin pencil. Experiments, including some by the author, have clearly shown the continuous nature of the change in normal cylindrical tubes with flat electrodes. In general the sharpness of the pencil seems to be independent of tube geometry.

The luminosity of the cathode glow, and thus of the pencil ray, is due to ionization processes initiated by positive ions or neutral particles (collisions of the second kind). This is shown partly by the fact that the pencil passes through the canal without appreciable alteration in structure. The cathode glow is a region of positive space-charge. Laporte (1939) stated that its luminosity is due to excitation or ionization by electrons (collisions of the first kind), but was probably confusing the cathode glow with the so-called first cathode layer which, at least in neon in certain conditions, is a separate phenomenon (Holst and Osterhuis, 1923; Seeman and Orban, 1935; Druyvesteyn and Penning, 1937).

Güntherschulze and Keller (1931) have clearly described the relation between the first cathode layer and cathode glow. The former is found (Ar, He, Kr, N_2 and O_2) for low cathode falls, and is due, as stated above, to excitation by electrons emerging from the Aston dark space. At higher cathode falls (loc. cit.) the cathode glow appears and becomes strong in Ne, N_2, H_2 and Xe, but remains weak in Ar, He and O_2 even with high cathode falls.

Further, electrostatic deflection experiments (see § 7) on the canal rays showed clearly that a large part of the luminosity is due to the presence of neutrals. Measurements indicate that the positive-ion stream to the cathode is not limited in extent to the region of luminosity, although the concentration of ions is larger at the centre of the tube (shown by Faraday-cylinder measurements and the burning of the cathode). Further evidence for the predominating influence of neutral particles in the formation of the pencil glow is produced when the relative importance of ionization by, and neutralization by electron transfer (*Umladung*) of, protons is considered.

These reactions may be represented as follows:—

(*a*) Ionization by protons: $H_1^+ + H_2 \rightarrow H_1^+ + H_1^+ + H_1^- + e^-$

or $H_1^+ + H_2 \rightarrow H_1^+ + H_2^+ + e^-$.

(*b*) *Umladung* (change of charge) of protons:

$$\overset{\text{slow}}{H_1^+} + \overset{\text{fast}}{H_1} \rightarrow H_1 + H_1^+.$$

The protonic mean free path at 0·15 mm. pressure is approximately, taking 7 kv. protons for which the Umladung effect is a maximum (Bartels, 1932), 6 cm. for ionization and 0·15 cm. for Umladung (Gerthsen, 1930). In these conditions, neutralization by electron transfer is more important than the production of more free electrons by a normal ionization process. It is difficult to assess the relative magnitudes of these two effects in a particular practical case. For instance, it is necessary first to determine the effective length of discharge path between anode and cathode in order that the p.d. for each mean free path may be found. This is difficult for the tubes described below (§ 6). If we assume an effective length of 3 cm. for 20 kv. discharge voltage, it is seen that the Umladung conditions given above (7 kv., m.f.p. = 0·15 cm.) are incompatible, although approximate corrections could be obtained by trial and error. It is hoped to discuss this problem further elsewhere.

The apparent diameter of the pencil glow in hydrogen in tubes with pressure ~0·1 mm. Hg and ~20 kv. operating potential for a current of ~10 ma. is approximately 2 mm., a surprisingly small figure. In order to explain this fully it would be necessary to determine the secondary process responsible for the glow, i.e. excitation of one of the products of the change of charge reaction, excitation or ionization by neutrals, recombination, etc. Bartels (1920) and Meyer (1937) stated that the cross-sections for ionization by neutral atomic hydrogen ions lie below the values given by Gerthsen (1930) for ionization by protons. See also the work of Yearian (1940).

The above discussion is not complete, but is intended to demonstrate the extreme importance of fast neutral particles in hydrogen glow discharges at pressures $\sim 0{\cdot}1$ mm. Hg.

Finally, the discharge mechanism as a whole should be outlined. (*a*) For high pressures (~ 1 mm. Hg) the discharge is maintained in stability by secondary emission from the cathode, the secondary electrons after acceleration producing a fresh supply of ions. (*b*) The pencil glow discharges can be maintained by ionization produced in the gas by the fast positive ions or neutrals.

Generally, therefore, for stability,

$$pn > 1$$

where for case (*a*) n is the number of ion pairs produced in the gas by one electron, ignoring diffusion losses, and p the number of secondary electrons produced at the cathode per incident positive ion. For case (*b*) n has the same meaning, but p is now the number of fresh electrons produced by each positive ion in the gas (Seeman and Orban, 1935, and Hailer, 1938). It is clear that secondary emission at the cathode for case (*b*) plays only a small part because a large proportion of the total positive current arriving at the cathode may be collected after passing through the canal.

In order to estimate the efficiency of this type of tube as an ion source, it is necessary to determine the total positive current arriving at the cathode by experiment and to express this as a fraction of the discharge current, read by a milliammeter in series with the tube. Such calculations are only made with great difficulty (Druyvesteyn and Penning, 1940, p. 136). Campan (1934) found for an almost normal cathode fall that the electronic current at the cathode was 6 to 16% for H_2 and O_2, but his conditions were different from those obtaining in the pencil-glow régime. In our case secondary emission from the cathode would, presumably, only be of importance because of the presence of large numbers of fast neutral particles. A high proportion of the ions passed through the cathode, and from the geometry of the tube (§ 6) we assume that photo-emission was small.

§ 4. MEASUREMENT OF HIGH-SPEED POSITIVE-ION CURRENTS

If a flat plate is exposed to a beam of positive ions with an energy of several kv., the resulting emission of secondary electrons can be shown, by the action of a suitable accelerating field, to be large enough to vitiate measurements of the primary beam. The problem has been investigated by many workers, but we shall refer only to the experiments of Hill and collaborators (1939) on H_1^+, H_2^+ and He_4^+ ions. They used surfaces which had not been specially prepared, so the results, a selection of which is given in table 1, are of particular interest with reference to continuously evacuated apparatus where baking is not practicable. For the following results, however, stable targets were obtained by continued degassing.

Table 1

Ion species	Target metal	Energy (kv.)						
		78	107	142	213	284	355	426
		Yield of secondaries/primary						
H_1^+	Cu	3·88	3·61	3·41	2·90	2·52	2·44	2·21
H_2^+	Cu	6·68	6·64	6·45	6·26	6·30	5·71	5·60

The results for Mo, Al and Pb were very similar, indicating that surface films were exerting a controlling influence. The secondary particles were very slow, e.g. for 213 kv. protons on Mo the great majority (~95 %) of secondaries could be turned back to the target by a retarding potential of 40 v. The increased emission with molecular ions is due to the latter splitting up into atomic ions (this effect is mentioned in § 2). In our case, an extra complication arises due to the proximity of the Faraday cylinder to the cathode (no focussing was used)

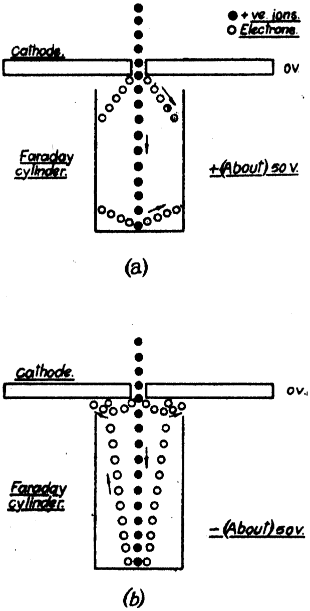


Figure 1. Effect of secondary emission of electrons on the measurement of positive-ion currents.

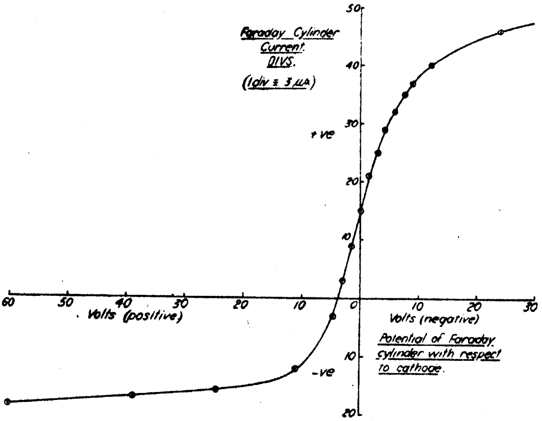


Figure 2. Effect of secondary emission of electrons on the measurement of positive-ion currents.

because of the secondaries arising in the canal. Consequently, a simple retarding field could not be used because the collector would receive negative particles from the canal. The effect is illustrated in figure 1. The characteristic of a simple Faraday cylinder, 7·5 cm. long × 2·5 cm. internal diameter, situated 3·7 cm. from the cathode, is shown in figure 2.

The problem was solved (see also Hailer, 1938) by the use of an intermediate electrode, maintained 120 v. negative with respect to the Faraday cylinder and 240 v. negative with respect to the cathode. The latter was, therefore, 120 v. positive with respect to the Faraday cylinder, but the smallness of the aperture in the intermediate electrode prevented field penetration. Emission of secondary electrons from the diaphragm into the Faraday cylinder because of bombardment by stray ions was very small. The final dimensions were:—Faraday (inner) cylinder 1 in. internal diameter, $^1/_{32}$ in. thick, 3 in. long; intermediate diaphragm $1^1/_2$ in. diameter, $^1/_{16}$ in. thick, aperture 1·1 cm. diameter spaced $^1/_8$ in. from the upper end of the inner cylinder, which, in turn, was spaced about 1 cm. from the cathode. The characteristics for this arrangement are shown in figure 3 ($\sim$10 kv. discharge). The outer cylinder served merely as a support for the diaphragm; both were at the same potential, and insulated from the inner cylinder.

Finally, mention should again be made of the neutral particles, which will not be affected by electric fields, and which may, in some cases, invalidate experiments such as those described above. We have ignored the existence of massive negative ions, which were later shown to be present only in very small numbers (§ 8). If it is desired to simplify ion tubes or similar devices in which collectors are required, other methods of eliminating spurious effects due to secondary emission are available, e.g. magnetic fields or the use of long narrow cylinders. In the latter case it may be taken as a rule that the field falls off to <5 % of its external value at a distance inside the collector equal to its diameter. These methods were used in the present research and found to be satisfactory. All the measurements given here, however, were taken with the composite cylinder.

§ 5. EXPERIMENTS WITH THE FIRST SOURCE

The methods of measurement outlined in § 4 were applied to a source of simple design (figure 4). The anode and cathode were water-cooled, the former being insulated from earth for 50 kv., with an isolated water system. The anode (B) and cathode were made of chromium-plated copper, to minimize field currents at the cathode and to give good thermal conduction. The flanges (C) were lapped and provided with centring screws (not shown). The cylinder (A), forming part of the cathode, was necessary to prevent waste discharges between C and the upper parts of B. The insulating cylinder was made of Pyrex, and the plate (D) of soda glass. Several cylinders (A) were made, each with a corresponding Pyrex cylinder, to give a constant {anode rim/cathode} spacing. Several lengths of anode chamber were tried.

Hydrogen was admitted through a small hole in the cathode (not shown in figure 4) from a needle valve and reservoir. No difference was found between tank hydrogen (99·5 to 99·7 % hydrogen, most of the impurity being oxygen) and that diffused through a palladium leak. This is hardly surprising, in view of the sealing technique adopted, i.e. the use of bitumen or Apiezon Q compound joints. The source gas pressures were measured with G.E.C. Pirani gauges, individually calibrated against a Macleod gauge and sealed on to a pipe soldered into the cathode block.

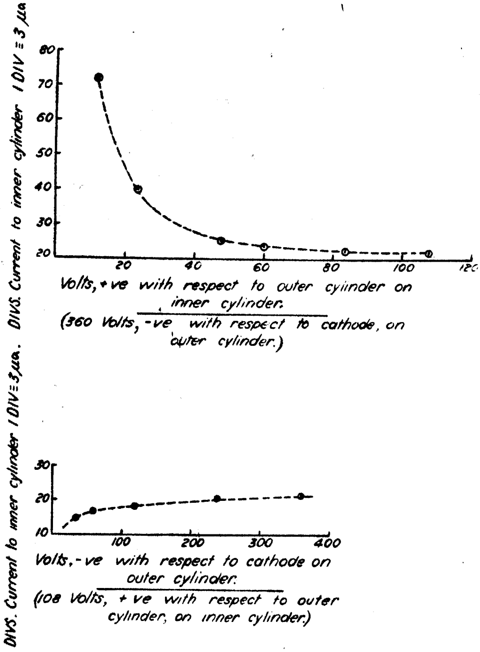


Figure 3. Experiments on composite Faraday cylinder.

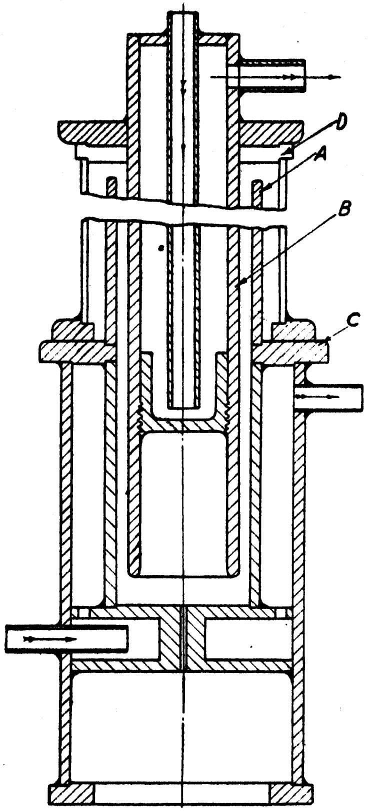


Figure 4. One-half size.

The pumping system (Metropolitan-Vickers oil-diffusion pumps) had a speed of about 20 l./sec. for air. The canal was 2 mm. diameter × 6·5 mm. long, these dimensions having been found by experiment to give the best yields for a reasonable gas flow through the tube.

A selection of results is given in figure 5 for an anode chamber 5 cm. long. The voltages ranged from 22 to 25 kv. for 0·092 mm. Hg. pressure and from 18·9 to 23 kv. for the higher pressure run.

Increasing the anode chamber to 10 cm. length gave slightly lower yields. Some results are given below:—

Table 2

5 cm. anode chamber				10 cm. anode chamber			
Hydrogen pressure (mm. Hg)	Tube voltage (kv.)	Tube current (ma.)	Canal-ray current (μa.)	Hydrogen pressure (mm. Hg)	Tube voltage (kv.)	Tube current (ma.)	Canal-ray current (μa.)
$0{\cdot}09_2$	22·2	2·1	60	$0{\cdot}08_8$	26·1	1	21
	23·4	2·8	75	,,	26·7	1·8	39
	24·2	3·6	90	,,	28	2·7	53
	24·8	4·4	105	,,	28·5	3·7	78
$0{\cdot}09_9$	16·2	1·6	40	0·11	13	1·2	25
	17·4	2·8	63	,,	13·4	1·8	33
	18·4	3·9	78	,,	13·5	2·3	40
	17·1	4·8	90	,,	13·6	2·8	42

A few readings at 15 cm. chamber length indicated that the yields were higher than those with a 10 cm. chamber and almost equal to those obtained with the 5 cm. chamber. Inserting a diaphragm with a 13 mm. central aperture in the lower end of the anode chamber increased the initial impedance, the required voltage for a given current and pressure being approximately doubled with reference to the data of figure 6—for instance, at 0·18 mm. Hg pressure 27·3 kv. (3·9 ma.) gave 141 μa. and 23 kv. (2·1 ma.) gave 96 μa.

The results may be briefly summarized. The yield increases steeply with main current but not with voltage. This indicates that the effective beam diameter does not increase rapidly with current. The cathode was usually burnt round the canal, indicating a beam diameter of $\sim$4 mm. Increasing the anode/cathode clearance from 2 mm. to 8·5 mm. made little difference to the performance of the tube—in fact it seemed that the beam diameter depended very little on electrode geometry for a tube without diaphragm, but no tests were made on very large tubes such as those used by Oliphant (1933), or Bouwers, Heyn and Kuntke (1937). The latter, working at pressures of 0·02 to 0·03 mm. Hg, obtained yields of 100 to 200 μa. at 25 kv. for discharge currents of 5 to 40 ma. respectively. Their data extend to 50 kv., 40 ma., at 0·021 mm. Hg pressure when the canal-ray yield was 600 μa.

It is of interest to mention that very small amounts of air mixed with the hydrogen caused the striking potential of the tube to be lowered considerably. This indicates that such sources might be operated at much lower voltages by the addition of an adulterant (e.g. nitrogen) without appreciably altering the proton content of the beam.

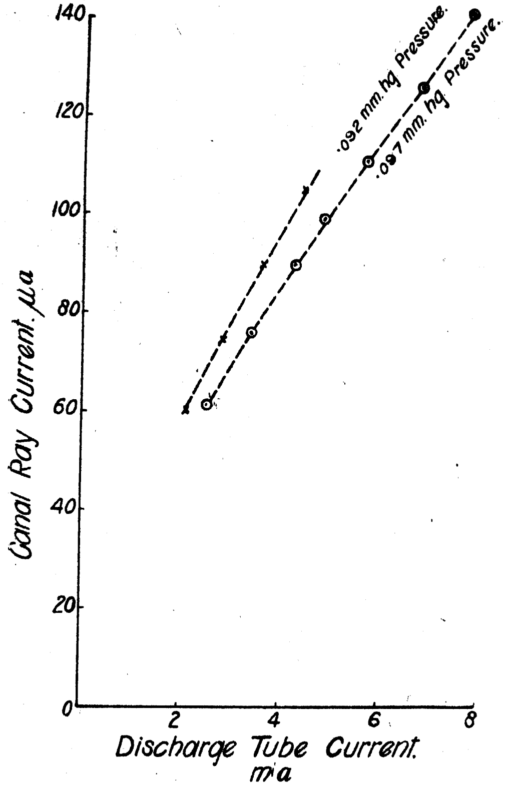


Figure 5.

Figure 5. Variation of canal-ray current with discharge-tube current (first tube).

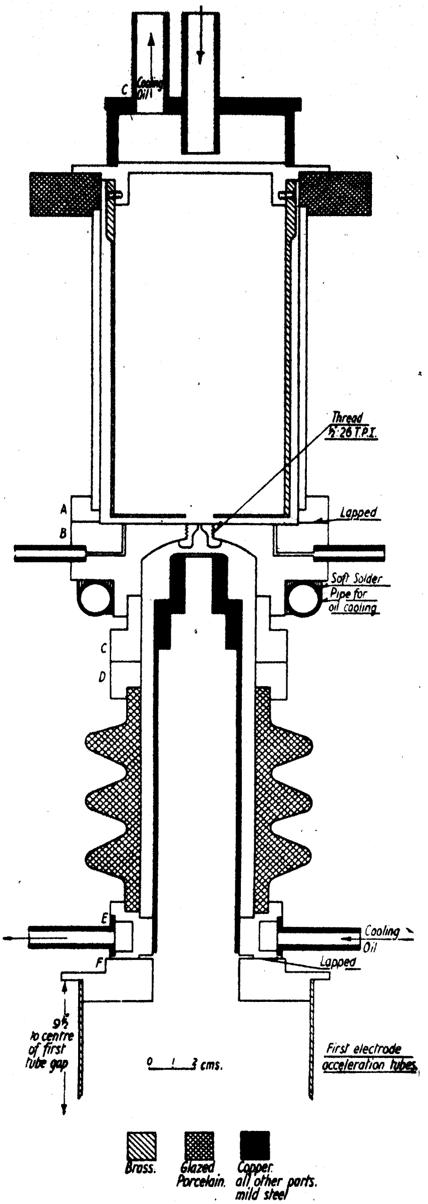


Figure 6.

Figure 6. The ion source and focussing electrode assembly. To avoid complicating the diagram, the following details are omitted: (1) Bakelite clamping bolts and nuts for ion source with brackets on B and C; (2) centring screws bearing on A with brackets on B; (3) centring screws bearing on C with brackets on D; (4) bakelite clamping bolts with brackets on B and F; (5) centring screws bearing on E with brackets on F; (6) steel clamping bolts with brackets on F and first plate of acceleration tube. Interior surfaces of ion source are polished.

§ 6. EXPERIMENTS WITH THE FINAL SOURCE

Many variations of the first source were tried, and simplifications resulted, a representative design being shown, with focussing assembly,* in figure 6. The tube diameter was increased in the hope that the operating pressure would be thereby reduced. The cathode cylinder was finally mounted on a lapped flange for screw centring (as in figure 4), the tube of figure 6 being centred mechanically. The anode insulation consisted of a porcelain ring (external diameter 6 in., thickness ¾ in., internal diameter 3½ in.), a ground recess being provided for the cathode cylinder. The tube length was 6 in., diameter of anode 3 in. The anode diaphragm aperture, the only critical dimension, was finally 13 mm. diameter, spaced 3 mm. from the cathode. The anode cathode side clearance was about 5 mm. With minor alterations, figure 6 represents the final design of tube.

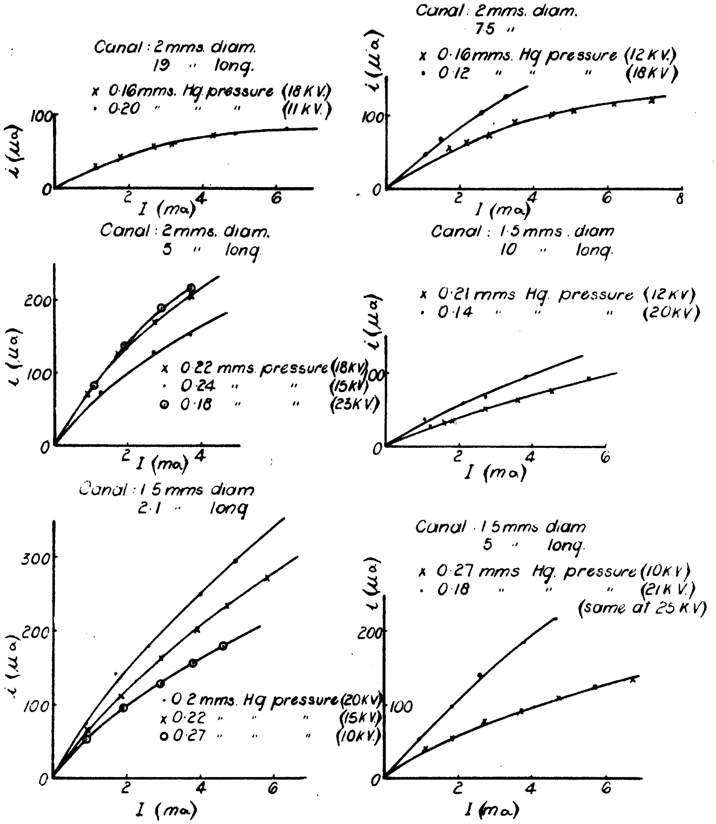


Figure 7. Canal-ray currents (i) plotted as a function of discharge current (I) for the type of tube shown in figure 6. Faraday cylinder 2·5 cm. from exit end of canal.

---

* The problem of focussing ion-beams will be the subject of a further paper.

It will be seen that the construction of canal-ray sources presents no mechanical difficulties; the centring is critical, but lining up is practicable mechanically.

The rectifier set (50 kv. maximum) for supplying power to the tube was provided with a Philips Metalix valve (Type 28202) for experiments with currents up to 30 ma.

It was necessary to age the tube for periods of almost an hour for stability at currents above 5 ma., but even prolonged treatment with many different forms of anode/cathode insulation showed that 35 kv. was the technical upper limit of voltage for our tubes. Attempts to increase the internal spark-over distances invariably lead to the establishment of stray discharges.

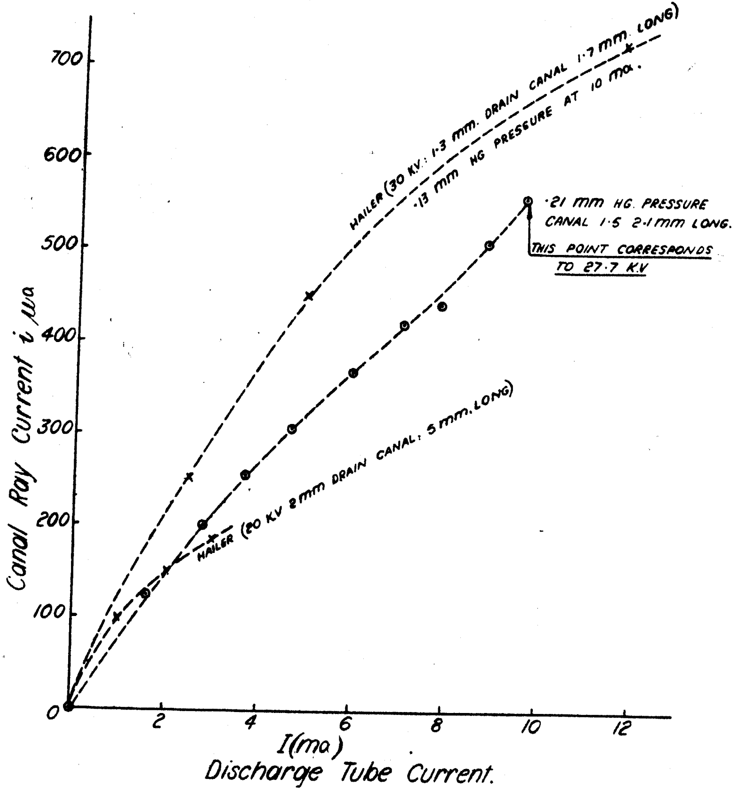


Figure 8. Variation of canal-ray current with discharge-tube current (second discharge tube).

Preliminary tests of this source with low currents are shown in figure 7. The purpose of this work was to establish the most favourable form of canal. It seemed that, although the gas consumption was high, we should concentrate on the 2 mm. diameter × 5 mm. long and 1·5 × 2·1 mm. long (approximately same pumping speed) canals. Figure 8 shows a more detailed curve for the latter, and includes some of Hailer's results (1938). Figures 9 *a* and *b* extend the data to higher currents, with the Faraday cylinder 2·5 cm. from the canal exit. Finally, figures 9 *c*, *d* and *e*, with a {cylinder/cathode exit} clearance of 1·2 cm., give the highest yields so far obtained with our tube. The differences in curvature between figures 9 *a* and *b* and figures 9 *c*, *d* and *e* were confirmed several times. It is hoped to give more data on this subject in a later publication. Figure 9 was taken with a 13-mm. diameter aperture in the anode diaphragm. In order to give as high a proportion of atomic ions as possible, the importance of which has been stressed above, a short canal is not beneficial, because the pressure gradient is high and the ion beam has little chance of losing intensity because of the ionization and Umladung processes. The primary particles are H_2^+; hence the probability of dissociation leading to the formation of H_1^+

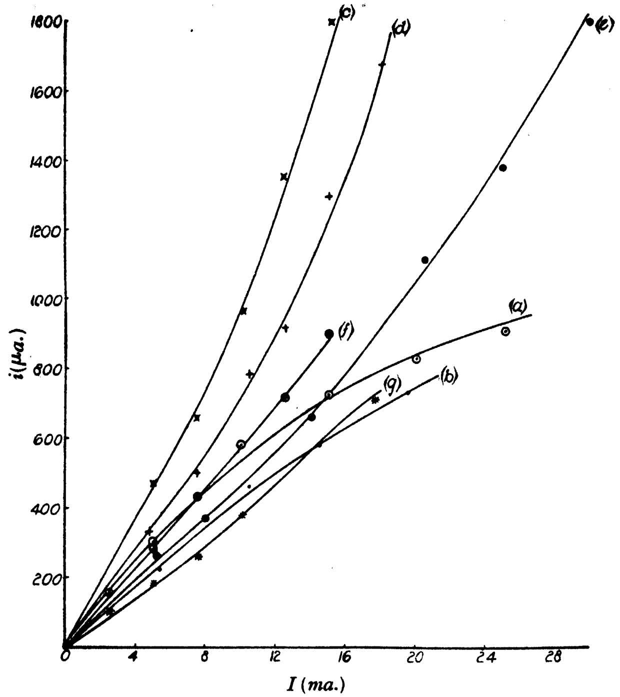


Figure 9. • Canal 2 mm. diam.: 5 mm. long. 0·16 mm. Hg. pressure (25 kv.).
⊙ „ 1·5 „ : 2·1 „ $0{\cdot}15_5$ „ „ (26 kv.).
× „ „ „ : „ „ 0·15 „ „
+ „ „ „ : „ „ 0·17 „ „
● „ „ „ : „ „ 0·18 „ „
⊕ Nitrogen: canal 1·5 mm. diam.: 2·1 mm. long. 0·06 mm. Hg pressure 25 kv.
✱ „ : „ „ „ : „ „ $0{\cdot}06_5$ „ „ 15 kv.

ions is small. On the other hand, with a long canal, the beam emerging from the discharge will traverse in the canal a longer path in gas at higher pressures, so that the opportunity for dissociation will increase but the emergent-beam intensity will decrease. A broad and long canal should, therefore, give the greatest yield with a low proton content (Schütze, 1939). For a given beam-strength the gas flow should always be arranged to be a minimum.

The fractional yield in general falls off at higher currents, presumably because of space-charge effects in the source (see, however, figures 9 *c*, *d* and *e*, which are unusual in this respect). It is believed that if the voltage were raised to 50 kv. the yields at ~20 ma. would be much increased.

It seems beneficial in many respects to take curves such as figure 9 at constant pressure and varying voltages rather than at constant voltage and varying pressure. It is always advisable to plot the current as one of the variables.

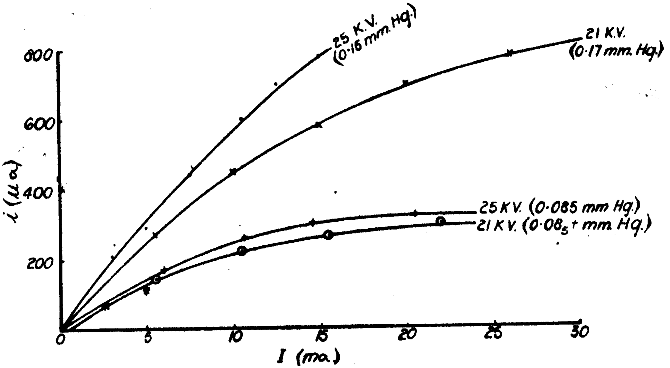


Figure 10. Effect of varying anode diaphragm aperture in canal-ray tube. Voltages and pressures given on curves.

13 mm. aperture in anode diaphragm:	• 25 kv.
	× 21 kv.
25·4 mm. „ „ „	+ 25 kv.
	⊙ 21 kv.
No diaphragm	✱ 30 kv.

The principal reason for the choice of plotting i/I was that the Pirani gauges, despite constant tapping, were prone to give variable readings, due largely to zero shifts (contact resistances, etc.). It was found that for a given current at a given voltage the yield was approximately repeatable from day to day even though the Pirani readings showed considerable variations. This was not considered a serious disadvantage, as the yield figures were meant to be relative only, provided that the pressures were known approximately, as the gas consumption of the sources is an important quantity. The latter is, unfortunately, high for tubes of the cold-cathode type with diaphragm.

Removal of the focussing diaphragm gave lower striking voltages. For instance, at 0·05 mm. Hg. hydrogen, 10·5 kv. sufficed to give 2 ma. (cf. figure 7). The yields obtained were low, e.g. 6 ma., 17 kv., 55 μa. and 5·5 ma., 33 kv., 115 μa. (pressures ~0·05 mm. Hg). These figures agree fairly well with those of Bouwers, Heyn and Kuntke (1937), also obtained without focussing diaphragm. Figure 10 shows the effect of increasing the size of the aperture in the diaphragm in our tube, and includes two points taken without a diaphragm. The reduction in striking voltage for increasing aperture size is to be noted. For all these tests the {anode diaphragm/cathode} clearance was about 4 mm.; the Faraday cylinder for figure 10 was 1·0 cm. from the exit end of the canal.

Voltage/current curves for the tube are shown in figure 11. Some yield-figures for tank nitrogen (~99·5 % pure) are added (figure 9) to show the differences in striking potentials (effect of increased efficiency of ionization, see Zwicky (1929)). The improved performance with small-diaphragm apertures may be due to increased pressure effects (" gas focussing ") or better electron-optical focussing or both. The anode diaphragm and cathode may be considered (Myers, 1939) as a pinhole lens, the focal length (f) of which may be expressed as

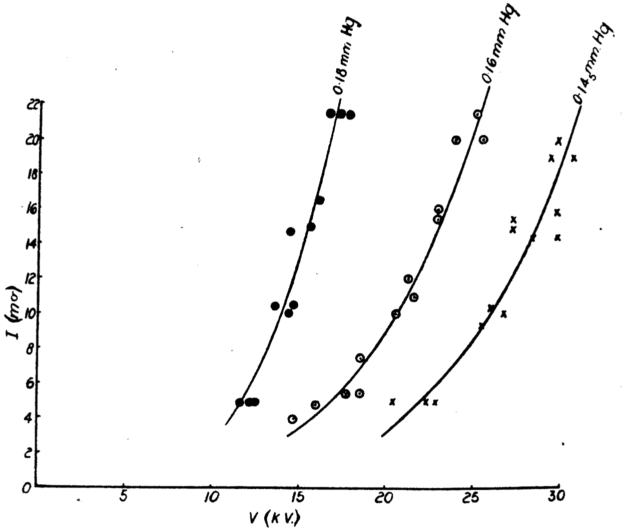


Figure 11. Current/voltage relation for canal-ray tubes 13 mm. anode diaphragm. Pressures (hydrogen) marked on curves.

$$\frac{1}{f}=\frac{V_2'-V_1'}{4V_g},$$

where V_g is the axial potential at the diaphragm, and V_2' and V_1' the potential gradients in the image and object space respectively. It should be noted that such lenses are not comparable with optical pinhole lenses, which do not have sharply defined focal lengths. Put $V_1'=0$, $V_2'=\frac{20\times10^3}{0\cdot4}$ v./cm., since the {anode diaphragm/cathode} clearance is 0·4 cm. and 20 kv. a representative tube voltage.

Since the treatment is clearly approximate, assume an infinite object distance (in fact ~10 to 20 cm.), so that f gives the image distance, which is required to be about 0·4 cm.; hence the beam is focussed on to the canal. V_g is thus found to be about 5 kv., which seems reasonable according to the published data on field plots. Our pinhole, of course, is of finite size, and errors are thereby introduced. However, we may assume that the focussing in our tubes at about 20 kv. is satisfactory, but the probability of a " cross-over " in the canal with subsequent beam divergence should not be overlooked. This is, perhaps, the most serious defect of sources of this type, as post-cathode focussing is rendered more difficult. It must be emphasized that a high beam current is required usually at the target of some high-voltage tube, probably at 6 to 12 ft. from the canal. Consequently, to obtain a high-current canal ray is only part of a formidable problem.

Finally, we should discuss canal sizes, and their bearing on tube yields, from the point of view of gas consumption and acceleration tube pressure. Consider the latter: most workers quote tube pressures of 2 to 5×10^{-5} mm. Hg., but Bouwers, Heyn and Kuntke (1937) operated a tube satisfactorily at 8×10^{-4} mm. Hg. Parkinson *et al.* (1938) stated that defocussing took place if the tube pressure exceeded $1{\cdot}5\times10^{-5}$ mm. Hg., in conflict with the findings of most workers. Assuming an acceleration-tube pumping speed of 300 l./sec., which is easily obtainable, an acceleration-tube pressure of 5×10^{-4} mm. Hg., and a source pressure of 0·2 mm. Hg., the canal pumping-speed must not exceed 750 c.c./sec. When using expensive gases, e.g. deuterium, it is, of course, necessary to reduce gas consumption as much as possible. The pumping system could be modified to circulate the gas, but this is generally not considered practicable. Using Knudsen's formula (Knudsen, 1909) for the pumping speed of narrow tubes (length $\gg$ radius; mean free path $\ll$ geometrical dimensions) we obtain 380 c.c./sec. for a 1·5 mm. diameter canal 2·1 mm. long, so that a reasonable safety factor is obtained. From a more accurate formula (Clausing, 1932), which allows for dimensions ~ mean free paths, the calculated speed was about 350 c.c./sec. (within ±5%). We may take the pumping speed S approximately as

$$S=k_1D^3/l,$$

where D = canal diameter, l = canal length, and k_1 is a constant. The output current varies approximately as $1/l$ (figure 7) until very short lengths are reached, and is independent of D, provided the canal area is greater than the focal spot, or the area (a) of the beam at the cathode. If $\pi D^2/4 \ll a$, the beam will be approximately proportional to D^2 for constant beam-current density. In practice it is found that $D=1{\cdot}5$ mm. is sufficiently large (see figure 7) for the beam current (i) to be given approximately by $i=k_2/l$ for varying D and l. Hence in general it is better to decrease D in order to reduce the pumping speed. This tends (see above) to increase the proton content of the beam.

§ 7. ELECTROSTATIC DEFLECTION OF CANAL RAYS

In the absence of detailed knowledge of ionization processes in the tube it it was considered worth while to investigate the possibility of electrostatic separation of the ion species. It is known, however, that for a given charge, the small angular deflection, θ, suffered by a charged particle (mass m, velocity v) travelling through an electrostatic field is proportional to

$$1/mv^2.$$

Consequently, if all ions (H_3^+, H_2^+, H_1^+) were to be formed at the same point in the discharge space, i.e. fall through the same potential difference, mv^2 would be a constant and there would be no separation. It was not known with certainty that this mechanism would obtain, so in view of other useful information that could be gained, a few tests were made.

Two deflector plates (5 cm. long × 6 cm. wide, with a separation of 2 cm.) were mounted in a horizontal pump-manifold 1 cm. behind the canal and observed through a glass plate in the end of the chamber.

It was found that perhaps 70 to 80 % of the luminosity (" hydrogen red ") of the canal rays was due to neutral particles, as the application of even 50 kv. failed to give detectable deflection of the main beam. The faint deflected beam was moved through 4·5 cm. at 25 cm. from the centre of the deflector system for 640 v. applied to the plates (beam positive), and gave a faint fluorescent patch on the glass plate. There was no visible separation of ion species; this was confirmed by Faraday-cylinder measurements. About 15 % of the no-field current was obtained when a large deflecting potential was applied to the plates, due possibly to liberation of some negative secondaries from the collector struck by fast neutral hydrogen atoms, but probably to neutrals gaining a positive charge by a reverse Umladung process after leaving the deflecting field.

The only point of interest in this experiment was the large proportion of neutral particles found in the emergent canal-ray beam. The estimate was, of course, entirely visual, and only has meaning if we assume the neutral and charged beams to be equally efficient in causing emission of light.

§ 8. BEAM ANALYSIS BY ELECTROMAGNETIC DEFLECTION

The importance of a high proportion of atomic ions in the beam has been stressed in § 2. Analysis by electromagnetic deflection only is sufficient (momentum resolution) if the velocity range of each group of particles at the target of the analyser is not high.

Some photographs of the spectrum obtained by allowing the analysed beam to strike a fluorescent screen were taken. It was not possible to perform tests on the ion source alone, owing to lack of time, and all the results given here refer to a high-voltage (~400 kv.) beam analysed by an electromagnet and suitable slit system at the earthed end of an acceleration tube.

The slit width was 3 to 4×10^{-3} in., ½ in. long, cut in a sheet of copper 10×10^{-3} in. thick. The slit was mounted in a flange at the entrance to a Pyrex tube, about 6 cm. diameter, which passed between, and extended 12 cm. beyond, the 4-in. pole pieces. The total distance from screen to slit was 23 cm. A beam current of ~1 μa. was sufficient to give reasonable contrast, higher currents increasing the background illumination too much for satisfactory measurement.

The proton line was identified first by measurement, and the identity of other lines deduced from a knowledge of their deflections (see table 3) taken from the fluorescent screen.

H_{2-1}^+ ions are those which enter the tube as H_2^+ ions and which split up in the acceleration tubes (eight gaps, so that eight ill-defined lines would be expected in the region 6–7 of table 3). If the field gradient were constant along the tube, and not concentrated at the focussing gaps, a continuous range of energies for the H_{2-1}^+ ions should be obtained. Some of the lines in table 3 are not visible in the photographs, so sketches of the spectrum are given in figure 12.

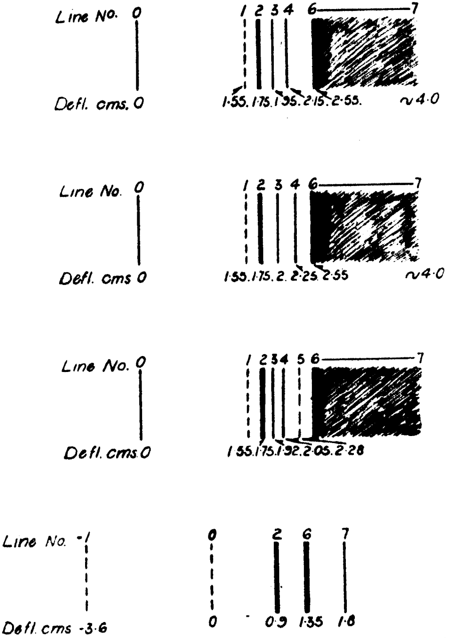
Table 3

Line number	Assumed identity of ions	Deflections (cm.)	
		Calculated	Measured
0	neutrals	0	0
1	H_3^+	1·65	1·65
2	H_2^+	2·0	2·0
3	H_2 or H_3	—	2·15
4		—	2·3
5	H_{3-2}^+	2·45	2·55
6	H_1^+	2·85 assumed	2·85
6–7	H_1^+ and H_{2-1}^+	—	—

The lines 3 and 4 were due to molecular ions formed in the tube; as the upper (high-potential) stages of the tube were shorted out, keeping the total voltage constant, and thus increasing the field-free region at the top of the tube, lines 3 and 4 merged into line 2. As the lower gaps were successively connected to earth, the lines 3 and 4 moved into line 6. For example, if a proton falls through half the accelerating-tube potential-difference before losing its charge to a neutral molecule (H_2) or attaching itself to an un-ionized atom, then the deflection for a given deflecting field will be [$\sqrt{(\frac{2}{3})}$ × proton deflection]. If this explanation of lines 3 and 4 is correct, it is surprising that there are only two or three lines in the spectrum for these processes. It is possible that they were due respectively to H_3^+ and H_2^+ (both formed in the accelerating tube). In the absence of more exact data, no definite conclusions may be drawn, but the reversed Umladung process should not be excluded as a possibility.

As mentioned above, we were not able to perform beam analysis on the ion source without acceleration tube. The heavier particles would probably be scattered more in the latter by space-charge effects than the atomic ions, though the current density, and, therefore, radial forces, were small in these experiments. In general, also, the effective absorption (due to all the collision processes) will be appreciably more for the molecular ions. The accelerating tube was operated at a pressure of ~10^{-4} mm. Hg., and was 8 ft. long. In order to gain information on the beam whilst minimizing the complicating effects of the tube, a few tests were made with one focussing gap (35 kv.) only in operation, the other electrodes being connected to earth. The spectra were too faint to be satisfactorily photographed, but taking the brightness of the H_1^+ line as 10, that of the H_2^+, H_2^+ and neutral lines respectively were estimated visually as 7–8, 2–3 and 3–4. Some negative ions were present in very small numbers; for the high-voltage work none were observed.

x displacements to scale, ×2 *Pole separation* 6·5 *cm. Source* ~9 *kv.* ~0·5 *ma.*
Beam ~1–5 *μa. Tube p.d.* 445 *kv.*



Two bottom gaps shorted out. Line 5 *moved into* 6. 2·4 *a magnet current. Shorting out top gaps makes lines* 3 *and* 4 *move into* 2.

Three bottom gaps shorted out. 2·4 *a magnet current.*

All gaps in operation. 2·4 *a magnet current.*

Top gap only (35 *kv. approx.*) 0·49 *a magnet current.* (*Note accuracy of* $\sqrt{2}$ *relation for deflections.*)

Figure 12. Analysis of hydrogen-ion beams in high-voltage acceleration tube.

Some quantitative measurements were taken with a Faraday cylinder, using larger deflections obtained with a special assembly (35° deflection, so that even with a ¼-in. slit width the resolution was adequate), figure 13 showing a typical curve. The calculated magnet field current for proton detection was 3·2 amp.; the $\sqrt{2}$ relation between the currents for H_2^+ and H_1^+ usually held only approximately, because the stray fields for this experiment were larger than for the photographic work described above, where the $\sqrt{2}$ factor was accurately obtained. The proton content was found from the curves (e.g. figure 13) to be about 40 to 50 %. No appreciable dependence of beam composition on source conditions was found, but it should be pointed out that the source only operates for a limited range of pressures, i.e. 0·08 to 0·25 mm. Hg.

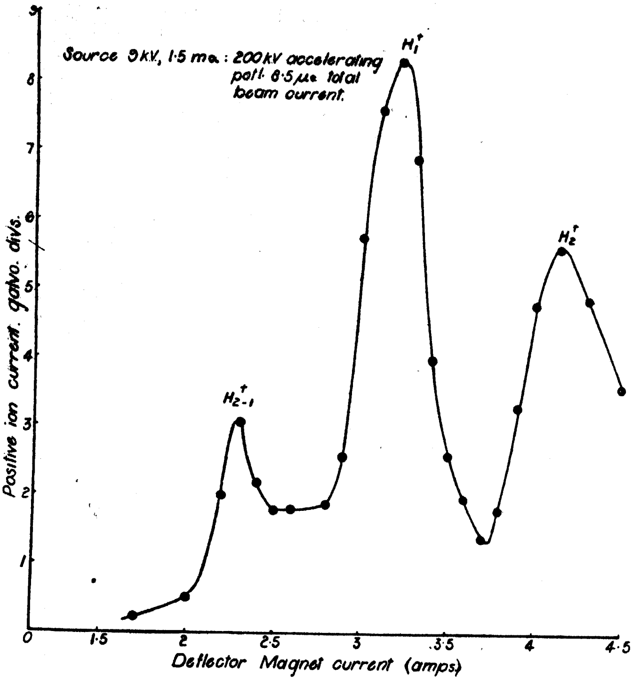


Figure 13. Analysis of hydrogen-ion beams in high-voltage acceleration tube (*continued*).

§ 9. CONCLUSIONS

A detailed study of the behaviour of canal-ray positive-ion sources for use with atomic disintegration apparatus gave the following results: (*a*) the total yield of positive ions could be made high (1 to 2 ma.); (*b*) the efficiency is low, though this usually is not of any importance; (*c*) the gas consumption is in general high; (*d*) the proportion of atomic ions is satisfactory, i.e. about 40 to 50 %.

This paper does not deal with the problem of focussing the canal-ray beam; given focussing voltages ~50 kv., this can be satisfactorily carried out. It is hoped to contribute another paper on this subject shortly.

§ 10. ACKNOWLEDGEMENTS

This research was carried out as part of a programme, on a high-voltage apparatus for atomic disintegration experiments, initiated and directed by Professor C. D. Ellis, F.R.S., whom the author wishes to thank for guidance and encouragement.

The author also wishes to express appreciation of the interest shown by Dr. T. E. Allibone, whose suggestions and advice have considerably helped the work.

Thanks are due to the Department of Scientific and Industrial Research for a grant covering part of the time during which the experiments were carried out, and to Dr. A. P. M. Fleming, Director and Manager, Research and Education Departments, Metropolitan-Vickers Electrical Co., Ltd., for permission to publish this paper.

REFERENCES

AMALDI, E., BOCCIARELLI, D., RASETTI, F. and TRABACCHI, G. C., 1939. *Ricerca Scientifica*, **10**, 623.
BARTELS, H., 1930. *Ann. Phys. Lpz.*, **6**, 957.
BARTELS, H., 1932. *Ann. Phys., Lpz.*, **13**, 373.
BOTHE, W. and GENTNER, W., 1937. *Z. Phys.* **104**, 685.
BOTHE, W. and GENTNER, W., 1939. *Z. Phys.* **112**, 45.
BOUWERS, A., HEYN, F. A. and KUNTKE, A., 1937. *Physica*, **4**, 153.
CAMPAN, T. I., 1934. *Z. Phys.* **91**, 111.
CHANG, W. Y., GOLDHABER, M. and SAGANE, R., 1937. *Nature, Lond.*, **139**, 962.
CLAUSING, P., 1932. *Ann. Phys., Lpz.*, **12**, 961.
DRUYVESTEYN, M. J. and PENNING, F. M., 1940. *Rev. Mod. Phys.* **12**, 87.
FRANZINI, T., 1938. *Atti Accad. Lincei*, **28**, 292.
GERTHSEN, C., 1930. *Ann. Phys., Lpz.*, **5**, 657.
GÜNTHERSCHULZE, A. and KELLER, F., 1931. *Z. Phys.* **72**, 143.
HAILER, C., 1938. *Wiss. Ver. Siemens-Werken*, **17**, 321.
HILL, A. G., BUECHNER, W. W., CLARK, J. S. and FISK, J. B., 1939. *Phys. Rev.* **55**, 463.
HOLST, G. and OOSTERHUIS, E., 1923. *Phil. Mag.* **46**, 1117.
KLEMPERER, O., 1939. *Electron Optics* (Cambridge: The University Press), p. 19.
KNUDSEN, M., 1909. *Ann. Phys., Lpz.*, **28**, 75.
LAMAR, E. S., BUECHNER, W. N. and COMPTON, K. T., 1937. *Phys. Rev.* **51**, 936.
LAMAR, E. S., SAMSON, E. W. and COMPTON, K. T., 1935. *Phys. Rev.* **48**, 886.
LAPORTE, M., 1939. *Décharge Electrique dans les Gaz* (Paris: Colin), p. 156.
MEYER, H., 1937. *Ann. Phys., Lpz.*, **30**, 635.
MYERS, L. M., 1939. *Electron Optics* (London: Chapman and Hall), p. 166.
OLIPHANT, M. L. E., 1935. Int. Conf. Physics, London, 1934. *Rep. Progr. Phys.* **1**, 145.
OLIPHANT, M. L. E. and RUTHERFORD, E., 1933. *Proc. Roy. Soc.* A, **141**, 259.
PARKINSON, D. B., HERB, R. G., BERNET, E. J. and MCKIBBEN, J. L., 1938. *Phys. Rev.* **53**, 642.
ROBERTS, R. B., 1937. *Phys. Rev.* **51**, 810.
SCHÜTZE, W., 1938. *Wiss. Ver. Siemens-Werken*, **17**, 341.
SEEMANN, H. and ORBAN, G., 1935. *Ann. Phys., Lpz.*, **23**, 137.
TIMOSHENKO, G., 1938. *Rev. Sci. Instrum.* **9**, 187.
YEARIAN, H. J., 1940. *J. Chem. Phys.* **8**, 24.
ZINN, W. H., 1937. *Phys. Rev.* **52**, 655.
ZWICKY, F., 1928. *Proc. Nat. Acad. Sci., Wash.*, **18**, 314.

THE DETERMINATION OF THE OPTICAL PROPERTIES OF THICK MAGNETIC LENSES, AND THE APPLICATION OF THESE LENSES TO BETA-RAY SPECTROMETRY

By R. E. SIDAY, B.Sc.,
Birkbeck College, London

Communicated by Prof. J. D. Bernal, F.R.S. Received 24 *November* 1941

ABSTRACT. An examination is made of the validity of the usual mathematical expressions from which the focusing properties of axially symmetrical magnetic fields are deduced; and conditions are established for good image formation.

A method of ray tracing under these conditions is developed and applied to a particular thick lens consisting of a solenoid wound with multiple layers and shielded with soft iron. The focal lengths and the positions of the principal planes are determined in terms of a parameter f, which is the focal length of a thin magnetic lens having the same value of $\int H^2 dx$. The dependence of the true focal length and the separation of the principal planes on f is obtained in a generalized form applicable to the determination of these quantities for any thick lens wound as a solenoid with multiple layers.

A differential method of applying the above conditions for good image production is used for an examination of the merits of coils with different types of windings, and for the estimation of the errors arising for any particular coil.

The advantages derived in β-ray spectrometry from the use of thick magnetic lenses are discussed, and a method is proposed for increasing the resolving power of the single-lens spectrometer in which use is made of a large rotation of the line image formed of a line object which can be obtained by winding a lens of suitable length.

§ 1. THE THEORY OF THE MAGNETIC LENS

THE simple theory of the magnetic lens (see, for example, Myers' *Electron Optics*) assumes that an electron at any point in an axially symmetrical magnetic field experiences a rotation about the axis whose velocity is

$$\frac{d\phi}{dt} = \frac{1}{2} \cdot \frac{e}{m} \cdot H_x, \qquad \ldots\ldots(1)$$

where H_x is the axial component of the magnetic field at the point considered, and that this rotational velocity produces an acceleration towards the axis which causes the focusing action of the lens. The usual cylindrical co-ordinates r, ϕ and x are used. The image is oriented at an angle ϕ to the object, where

$$\phi = \frac{1}{2} \frac{e}{m} \int \frac{H_x}{v_x} dx, \qquad \ldots\ldots(2)$$

v_x being the axial component of the velocity of the electron, and the integration being taken along each ray.

It is convenient (and legitimate) to consider the focusing of the lens and the rotation of the image separately, in which case the focusing of the lens becomes strictly analogous to that of an optical lens.

If v_x is assumed to be constant, equation (1) with the other equation of motion leads to

$$\frac{d^2r}{dx^2} = -\frac{r}{4}\left(\frac{e}{m}\right)^2 H_x^2/v_x^2, \qquad \ldots\ldots(3)$$

which is the equation of the trajectory of the electron within the lens. For a thin lens Busch (1926 and 1927) calculated the focal length,

$$\frac{1}{f} = \frac{1}{4}\left(\frac{e}{m}\right)^2 \int_{-\infty}^{+\infty} \frac{H_x^2}{v_x^2}\,dx, \qquad \ldots\ldots(4)$$

by directly integrating equation (3), assuming that r was constant within the lens for any electron. The constancy of r is only true in general for a thin lens.

These expressions show, however, how to obtain the condition for the accurate focusing of any magnetic lens. In this paper, only rays that are not skew to the axis outside the lens are considered. Equation (1) is accurately true everywhere in an axially symmetrical field if the equation

$$\frac{1}{2}\cdot\frac{\partial H_x}{\partial x} = -\frac{H}{r} \qquad \ldots\ldots(a)$$

holds at all points of the field. Except in the case of an infinite uniform field, this condition is not strictly consistent with Laplace's equation, although at a point in an axially symmetrical field it is more nearly satisfied the closer this point is to the axis.

Equation (2) shows that $\int \frac{H_x}{v_x}\,dx$ must be the same for all rays if the image of an object is to be oriented with respect to the object without angular spread. Now, in order to obtain equation (3) from equation (1), it is necessary to assume that the axial component of the velocity of the electron is unchanged in its passage through the lens. This leads to the condition

$$v_x = \text{constant}, \qquad \ldots\ldots(b)$$

while the condition of formation of images oriented without angular spread therefore reduces to

$$\int H_x\,dx = \text{constant}, \qquad \ldots\ldots(c)$$

where the integration is taken along the trajectory.

The degree of accuracy with which the lens produces equal rotation for all beams depends on the degree of accuracy with which the above conditions (*a*), (*b*) and (*c*) can be satisfied at all points along the rays.

Identical total rotation for all rays originating from the same point is a condition for accurate focusing. Another condition follows under the same assumptions from (4), by integrating along the trajectory:—

$$\int H_x{}^2\,dx = \text{constant}, \qquad \ldots\ldots(d)$$

a condition independent, in general, of the constancy of $\int H_x\,dx$.

In geometrical optics, the defects of the image are exactly and elaborately classified, and this classification is of very great importance in the construction of lenses free from particular aberrations. It is not clear that such a procedure has the same importance for the magnetic lens. There is no method of combining lenses to reduce a particular aberration. In the design of a lens, emphasis should be placed on direct construction to the required standard of focusing rather than on the measurements or calculations of aberrations with a view to their subsequent correction. This can be effected by satisfying the four conditions above.

§ 2. THE THICK MAGNETIC LENS

Equation (3) shows that the *thin* magnetic lens will only focus to the degree of accuracy with which the above conditions are satisfied, and that equation (4) gives the focal length of the lens.

If the magnetic field spreads along the axis for a distance that is not small compared with the focal length, the lens is thick, r is no longer even approximately constant for an electron during its passage through the lens, so that while (2) still gives the rotation, and (3) still determines the paths of the rays when the above conditions hold, the focal length will no longer be given by (4).

In order to examine the rules for the focusing of a thick lens, such a lens was constructed as a solenoid about 20 cm. long, wound with multiple layers. The field along the axis was measured, and assuming that the above conditions were satisfied, the trajectories of electrons were traced through the lens, in order to determine its focal length, principal planes and its focusing properties.

§ 3. RAY TRACING THROUGH A THICK LENS

Rays were traced in terms of the parameter f defined by equation (4) above. In order to make the equation depend only on relative values of H, and to introduce f explicitly, equation (3) is written in the form

$$\frac{d^2r}{dx^2} = -\lambda\,.\,r\,kH_x{}^2, \qquad \ldots\ldots(5)$$

where k is an arbitrary constant arising from the measurement of H. So that now

$$\left(\frac{dr}{dx}\right)_2 - \left(\frac{dr}{dx}\right)_1 = -\frac{(r_1+r_2)}{2}\,.\,\lambda\,.\int kH_x{}^2dx, \qquad \ldots\ldots(6)$$

and if between two points on the axis of the lens

$$\int_1^2 kH_x{}^2dx = {}_2A_1, \qquad \ldots\ldots(7)$$

$$\left(\frac{dr}{dx}\right)_2 - \left(\frac{dr}{dx}\right)_1 = -\frac{(r_1+r_2)}{2}\cdot\frac{{}_2A_1}{fA}, \qquad \ldots\ldots(8)$$

where

$$A = \int_{-\infty}^{\infty} kH_x{}^2 dx. \qquad \ldots\ldots(9)$$

For purposes of ray tracing, the region over which the field extended was divided into 1 cm. steps, so that

$$r_2 = r_1 + \frac{1}{2}\left\{\left(\frac{\partial r}{\partial x}\right)_2 + \left(\frac{dr}{dx}\right)_1\right\}. \qquad \ldots\ldots(10)$$

$$\left(\frac{dr}{dx}\right)_1\left(1 - \frac{{}_2A_1}{4fA}\right) - \left(\frac{\partial r}{\partial x}\right)_2\left(1 + \frac{{}_2A_1}{4fA}\right) = r_1\cdot\frac{{}_2A_1}{fA}. \qquad \ldots\ldots(11)$$

Thus, knowing an initial r, and dr/dx, the rest of the ray can be traced. These initial points are taken outside the magnetic field. The f's appearing in these formulae are the focal lengths the lens would have if it were infinitely thin, and they can be directly chosen in cm. for purposes of ray tracing. ${}_2A_1/A$ for any interval will depend only on the geometrical form of the winding. By tracing suitable rays, the real focal lengths, and the positions of the principal planes corresponding to different values of f, were determined.

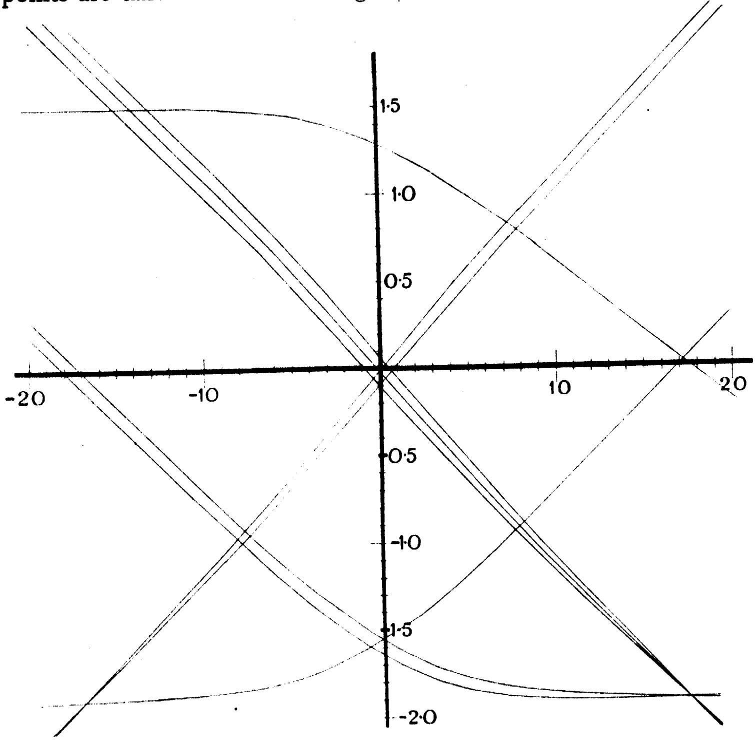


Figure 1.

This method of ray tracing gives the value of r at any point on the ray more accurately than it gives dr/dx at this point. The cardinal planes were therefore fixed by tracing a parallel bundle of rays in the object space inclined at some definite angle to the axis. These intersect in the image space on the image focal plane, and the ordinate of the intersection is related to the object focal length by the relation

$$r = F \, . \tan\theta, \qquad \ldots\ldots(12)$$

where F is the true focal length of the lens and θ the angle that the parallel bundle makes with the axis in the object space. (This is the most general definition of focal length in geometrical optics, being the only one applicable to skew rays). The alternative method, in which the ray traced is initially parallel to the axis in the object space, is not as reliable, because the principal planes and the focal length are determined from the final slope of this ray in the image space.

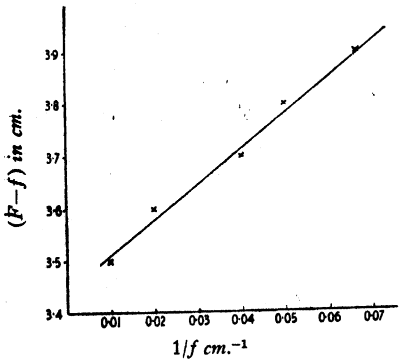


Figure 2.

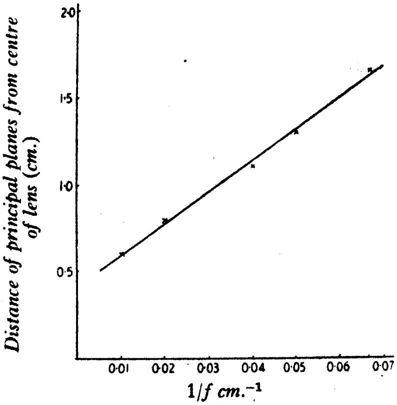


Figure 3.

For the lens described above, the distribution of the magnetic field along the axis was determined experimentally. The value of $f = 15$ cm. was chosen, and this example was fully traced with parallel bundles in the object and in the image spaces. These tracings are shown in figure 1. The actual focal length was found to be 18·9 cm. The principal planes were 1·65 cm. from the centre, and were " crossed ".

This fully-investigated case of $f = 15$ cm. showed that two parallel rays in the object space were sufficient to fix the cardinal points of the lens (since the cardinal planes are flat surfaces), and accordingly the lens was investigated for f equal to 20 cm., 25 cm., 50 cm. and 100 cm. The values of the corresponding focal lengths, and the positions of the principal planes, are given in table 1. Figure 2 shows how these may be plotted to give a straight-line graph. $(F-f)$ was plotted against $1/f$, and is seen to give a very good straight line. Also, figure 3 shows the distance of the principal planes from the centre of the lens plotted against $1/f$. The linear relationship is again satisfactory. Two interesting features of these curves are that the differences $(F-f)$ between the true

Table 1

f (cm.)	$(F-f)$, the difference between the true focal length and f (cm.)	Distance of the focal point from centre of lens (cm.)	Separation of principal planes from centre of lens (cm.)
15	3·9	17·5	1·65
20	3·8	22·5	1·3
25	3·7	27·6	1·1
50	3·6	52·8	0·8
100	3·5	102·9	0·6

focal lengths F and the parameter f are nearly constant over a wide range of values of f; and that even for very large focal lengths there remains a definite displacement of the principal planes from the centre of the lens.

§ 4. THE SHORT MULTI-LAYERED SOLENOID

The calculations above apply only to a particular thick lens. The form of the results as given in figures 2 and 3 is clearly more general. By ray-tracing methods, the determination of the focal lengths and positions of the principal planes was effected for single-layered solenoids of lengths 20 cm. and various radii. From these figures, given in table 2, it is seen that these lenses have remarkably similar properties, and that the focal lengths described in terms of the f's remain constant over a large range of radii. Because of this property, the curves shown are sufficient to fix the optical properties of a very large variety of lenses wound in the form of solenoids. For if the length, radius and the parameter f for any coil are altered in the same ratio, the value of $(F-f)$ will be altered in this ratio, as will the separation of the principal planes from the centre of the coil; and these quantities can be found from table 2. This will be particularly easy as

Table 2. For a lens wound as a single-layered solenoid, 20 cm. long

Radius (cm.)	$(F-f)$ (cm.)	Distance of principal plane from centre (cm.)
10	4·0	1·5
5	3·9	1·4
2·5	2·4	0·3

the properties of the lens are not sensitive to changes in radii, and the results will still give a good approximation even if the ratio of the radii is not accurate.

Again, the same property allows the effects of the multiple layers with which any practical coil would have to be wound to be taken into consideration. For unless the solenoid is wound with a very deep channel it will usually be sufficiently accurate to calculate f for the entire coil and to use the values of $(F-f)$, and the separation of the principal planes, for a mean layer of it.

§ 5. THE ACCURACY OF FOCUSING OF THE LENSES

The method of ray tracing described in § 3 was used above to determine focal lengths under conditions such that (*a*), (*b*), (*c*) and (*d*) were satisfied. Under these conditions it is clear that the " thickness " of the lens does not introduce any fresh aberrations (compared with a thin lens) of these types that can be described by tracing rays in the plane of the axis.

The formulae (10) and (11) used for ray tracing show that all rays initially parallel to the axis are brought to the same focus. Thus, under the specified conditions, the thick lens is free from spherical aberration. In fact, under these conditions, equation (3) becomes a linear homogeneous differential equation, and it follows that a point on the axis is accurately focused to a point on the axis. Again actual ray tracing of bundles both parallel to, and inclined at a definite angle to, the axis gives the same positions of the cardinal planes and the same values of the focal lengths.

According as the conditions (*a*), (*b*), (*c*) and (*d*) are not satisfied, inexactness is introduced into the focusing of the lens. It is not proposed to attempt to discuss here the aberrations produced by specific departures from these conditions, but in order to construct a lens of given quality it is necessary to see how to satisfy the conditions to the required degree of accuracy.

Conditions (*a*) and (*b*) must be satisfied for each point of a ray. Conditions (*c*) and (*d*) are integral conditions to be satisfied over the whole path of the ray. Thus, to determine the actual rotations of the different rays forming an image, (*c*) would have to be integrated along each ray. Alternatively, a differential method can be adopted to simplify the determination of the optical properties of a given magnetic field. Instead of considering, for example, the total rotation for a ray, the rate of rotation for an electron at each point of the field can be considered and the total rotation added up arithmetically. The above conditions for accurate image formation will now be considered in greater detail.

(a) *The constancy of v_x for all rays*

This condition fundamentally limits the available aperture for a permitted inaccuracy in the performance of any lens. Consider the formation of the image of a point on the axis which emits electrons of velocity v. The axial component of the velocity of an electron entering the lens inclined at an angle θ to the axis is $v_x = v \cos \theta$, compared with the value $v_x = v$ for an electron moving along the axis. Inside the lens, the ray will be bent round until at one point it has no radial component of velocity. Here, however, v_x will still be less than v, owing to the tangential velocity the electron has acquired. This tangential velocity is given by equation (1). Suppose that at any point (x, r) the axial component of the magnetic field is H_x and the electron is moving parallel to the axis. The tangential velocity here is $\frac{r}{2}\left(\frac{e}{m}\right)H_x$, so that the axial velocity v_x is given by

$$v_x^2 = v^2 - \frac{r^2}{4}\left(\frac{e}{m}\right)^2 H_x^2. \qquad \ldots\ldots(13)$$

Let

$$\int_{-\infty}^{\infty} H_x^2 dx = H_x^2 l, \qquad \ldots\ldots(14)$$

where the length l defined is related to the thickness of the lens.

Then,

$$v_x^2 = v^2\left[1 - \frac{r^2}{4}\left(\frac{e}{m}\right)^2 \frac{\int H_x^2 dx}{lv^2}\right] \qquad \ldots\ldots(15)$$

or

$$v_x = v\left[1 - \frac{r^2}{2f} \cdot \frac{1}{l}\right]. \qquad \ldots\ldots(16)$$

Before the electron entered the lens it follows from simple geometrical considerations that

$$v > v_x \geqslant v\left[1 - \frac{r^2}{2f} \cdot \frac{1}{f}\right], \qquad \ldots\ldots(17)$$

according to the position of the point object, the sign of equality holding for the object placed at the object focal point.

These equations exhibit the striking superiority of a thick lens over a thin lens in this respect.* For the latter, $f \gg l$, and the angular aperture for a permitted variation in v_x is much less than that resulting from the cosine factor of the velocities of the rays entering the lens, but by winding the lens of such a length that $f \simeq 1$, apertures up to the full one fixed by the cosine factor can be used according to the position of the point object. Since, in the symmetrical lens, the principal points lie near the centre of the coil, l cannot conveniently be much greater than f. Moreover, due to the continuous bending of the rays in a thick lens, the value of r in equation (16) is smaller for a thick lens than for a thin one.

(b) *The form of the field*

The other conditions, (*b*), (*c*) and (*d*), depend on the form of the field, and can conveniently be considered together. The axially symmetrical magnetic field can be defined in the usual way by the potential

$$\Omega = A_0 - \frac{r^2}{2^2} \cdot \frac{\partial^2 A_0}{\partial x^2} + \frac{r^4}{2^2 4^2} \cdot \frac{\partial^4 A_0}{\partial x^4}, \qquad \ldots\ldots(18)$$

where A_0 is a function of x only.

---

* Indeed, it appears that the thinner the lens, the smaller is the aperture of the cone of rays accurately focused. A perfectly thin lens would not focus.

The condition $\frac{1}{2}\,\partial H_x/\partial x = -H_r/r$ shows that equation (1) would be strictly true if the axial component H_x of the field was everywhere given by

$$H_x = -A_0' + \frac{r^2}{8}A_0''' - \dots \qquad \dots\dots(19)$$

instead of by

$$H_x = -A_0' + \frac{r^2}{4}A_0''' - \dots \qquad \dots\dots(20)$$

Of course, except in certain trivial cases, the field given by equation (19) cannot exist, as it is inconsistent with Laplace's equation, but (19) is a good approximation to (20) provided that

$$\frac{r^2}{8}\,|A_0'''| \ll |A_0'|, \qquad \dots\dots(21)$$

and the accuracy with which (1) holds depends on the degree of accuracy with which this condition is satisfied.

Conditions (*c*) and (*d*) are actually integral conditions. They can, however, be replaced by the much simpler condition that H_x should at any point be independent of the distance of this point from the axis. In this form, the condition is too stringent, but any field which satisfies the simplified condition will certainly satisfy (*c*) and (*d*). For the axially symmetric field the new condition becomes

$$\frac{r^2}{4}|A_0'''| \ll |A_0'|. \qquad \dots\dots(22)$$

As the condition derived from (*b*) is less stringent than this, and as departure from exact fulfilment of both conditions produces errors in the same direction, the latter can be used alone.

For the examination of actual magnetic fields, the differential method of estimating the errors produced from these conditions is far simpler to apply. As far as the rotation of the image is concerned, the percentage variation in the rate of this, due to the electron being at the point (x, r) instead of being very close to the axis, is

$$100\,\frac{H(x,r) - H(x,0)}{H(x,0)}$$

$$= 25\,r^2\left|\frac{A_0'''}{A_0'}\right|. \qquad \dots\dots(23)$$

This method is worked out in figures 4 to 7. If

$$25r^2|A_0'''/A_0'| \leqslant n,$$

a region is defined in the lens by the inequality:

$$r \leqslant 1/r\sqrt{n|A_0'''/A_0'|}, \qquad \dots\dots(24)$$

where the error in the rate of rotation is everywhere $\leqslant n\%$. Such contours

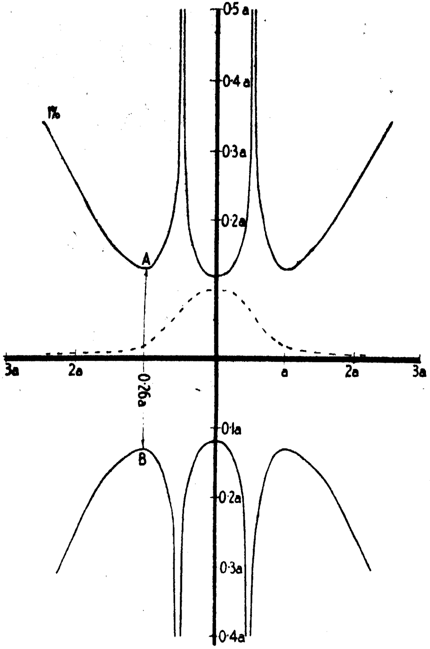


Figure 4. Circle of wire of radius *a*.

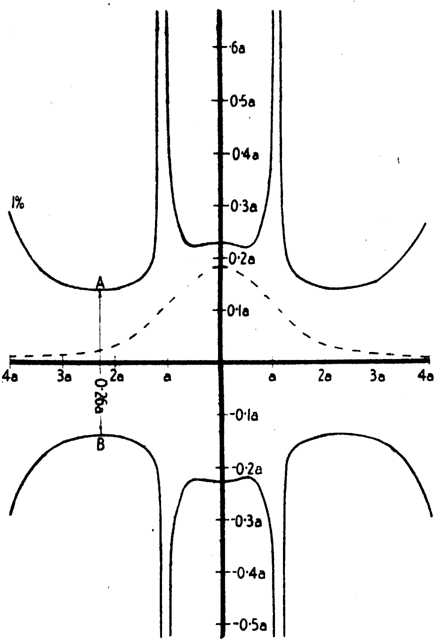


Figure 5. Solenoid of radius *a* and length *a*.

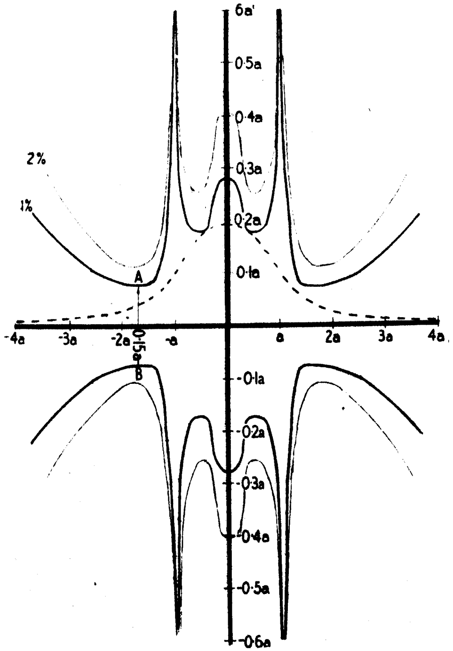


Figure 6. Solenoid of radius *a*/2 and length *a*.

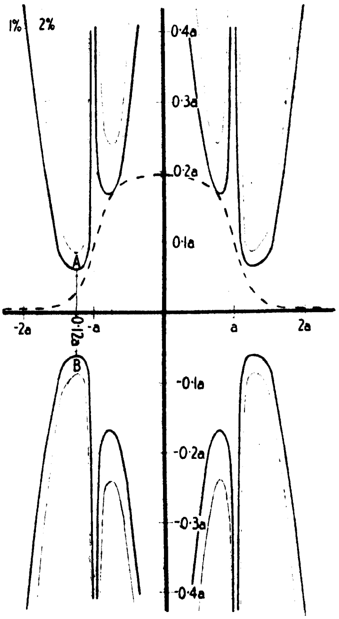


Figure 7. Solenoid of radius *a*/4 and length *a*.

The figures show the regions within which the electron tracks must be confined so that the deviations in their rates of rotation shall always be less than 1 %. The regions between the asymptotes produce deviations in the opposite directions to the regions outside them.

for errors of 1% are drawn for various fields. If the aberrationless ray tracks are drawn on top of these for the field which is to be used, the total variations in the rotation can be added up for each ray. The method is also useful for comparing different fields. The accompanying figures show that, although by winding a solenoid of small radius, the actual size of the entrance to the lens for a given permitted error is reduced compared with a solenoid wound with a larger radius, nevertheless this reduction is quite small. Further, any ray entering the lens passes through two regions where the errors in rotation are in opposite directions. The smaller the radius of the winding, the narrower and the closer these regions become. Further, they are pushed out into a position near the end of the coil where the actual field is weak, and an error in the rate of rotation contributes but little to the total rotation. Apart from the obvious advantages of cheapness, ease of winding, and gain in magnetic field derived from winding coils of small radii, the smaller the radius, the more rapidly does the field diminish outside the actual length of the winding of the solenoid. The merits of a particular winding can always be judged by drawing the aberrationless ray tracks on top of the error contours constructed for the winding as in figures 4 to 7, but it would appear that, except for special purposes, little advantage, if any, is to be derived from the use of solenoids of large radii.

§6. THE DESIGN OF A β-RAY SPECTROMETER

Attempts made hitherto by Klemperer (1938), Cosslett (1940) and others to use a magnetic lens as a β-ray spectrometer have all involved the use of thin lenses and the focusing of a point on the axis to form a point image by the method of focal isolation that was used years ago by R. W. Wood for his investigation of infra-red spectra. Recently Witcher (1941) has used a homogeneous magnetic field. To prevent unresolved rays along the axis from entering the receiver, they used a stop on the axis.

The obvious advantage to be derived from the use of thick lenses over thin lenses is that much weaker fields can be used. As Witcher points out, this becomes very important for the focusing of high-energy electrons. Thus to focus β rays of 15 Mev., Cosslett constructed a powerful electromagnet to give a focal length of about 40 cm. with a pole gap—i.e. lens thickness—of 5 cm. A solenoid wound for 50 cm. and used to give a focal length of 50 cm. would need a field along the axis of only 2000 gauss to focus these electrons. What is more important is that, according to the discussions of aberrations given here, such a lens should have far better properties.

The special feature of magnetic focusing is the rotation of the image with respect to the object. While for most of the applications of magnetic lenses this has only a nuisance value, it has been turned to specific use in a single-lens β-ray spectrometer constructed here in an attempt to increase the resolving power of these instruments.

Fundamentally, the method consists in forming a line image of a line source; the electrons are collected by a line receiver. The lens is designed to give a rotation of about 90°, and to be as free as possible from aberrations.

Very approximately for a solenoid, if

$$\left.\begin{aligned}\int H_x\,dx &= Hl,\\ \int H_x^2\,dx &= H^2l,\end{aligned}\right\} \qquad \ldots\ldots(25)$$

then

equations which are sufficiently true for purposes of design. Thus to get the right rotation

$$\frac{1}{2}\left(\frac{e}{m}\right)Hl/v=\pi/2,$$

which fixes the value of the product Hl. Now the approximate focal length is given by

$$\frac{1}{f}=\frac{1}{4}\left(\frac{e}{m}\right)^2\frac{H^2l}{v},$$

and the value of H is chosen here so that the value of f obtained is somewhat less than l; for, as shown above, the lens is most free from aberrations when the length of the winding is as great as possible compared with the focal length. The lens actually constructed here has a length of winding of about 20 cm., to be operated at an f of 15 cm.

ACKNOWLEDGEMENTS

I should like to thank Professor P. M. S. Blackett, F.R.S., for the encouragement that he gave at the beginning of this work, and Professor J. D. Bernal, F.R.S. for his continued support. I owe much to my former colleague Dr. W. Ehrenberg, now at Electric and Musical Industries, Ltd., for many useful and encouraging discussions.

REFERENCES

Bush, H., 1926. *Ann. Phys., Lpz.*, **81**, 976 ; 1927. *Arch. Elektrotech.* **18**, 583.
Cosslett, V. E., 1940. *Proc. Phys. Soc.* **52**, 511.
Klemperer, O., 1935. *Phil. Mag.* **20**, 545.
Witcher, C. M., 1941. *Phys. Rev.* **60**, 32.

THE DETERMINATION, BY SMOKE MANOMETER, OF THE RATIO OF DENSITY OF TOWN GAS TO THAT OF AIR AT THE SAME TEMPERATURE AND PRESSURE *

BY G. F. C. SEARLE, Sc.D., F.R.S.,
Cavendish Laboratory, Cambridge

Received 6 *March* 1942

ABSTRACT. The buoyancy effect of a column of coal gas is balanced against that of a column of hot air. At their upper ends, the columns are in free communication with the atmosphere. The lower ends of the tubes containing the gas and the hot air are joined to a smoke manometer; communication between the columns is established by turning a tap. The absence of drift of the smoke indicates that the pressures at the lower ends of the columns are equal. The effective height of the hot-air column is h cm., and the absolute temperature of the steam jacket surrounding it is θ'. The balancing height of the gas column is k cm.; the temperature of the gas is θ, the atmospheric temperature. Then, exactly,

$$\sigma/\rho = 1 - h(\theta' - \theta)/k\theta',$$

where the densities σ, ρ of gas and air are measured at a temperature and a pressure which are the same for both gas and air.

In a second method, the buoyancy effect of the gas is balanced against the defect of pressure due to the combination of a whirling arm and a static-head tube described in an earlier paper. If v be the velocity in its path of the test hole of the static-head tube and k be the height above the centre of the manometer of the upper (open) end of the gas column,

$$\sigma/\rho = 1 - v^2/2gk.$$

The value of σ/ρ may also be found by weighing an aluminium hot-water bottle (1) full of air and (2) full of gas, the temperature and pressure being the same in each case.

§ 1. THE SMOKE MANOMETER

THE manometer is a modification of that shown in figures 6, 7 in a paper —quoted later as "Paper A"—by G. F. C. Searle (1941). Its details are shown in figures 1, 2, 3 below. A rectangular block of brass $3\frac{1}{2}$ inches in length and 1×1 inch in section is used. An axial duct $\frac{1}{8}$ inch in diameter is bored from end to end. The taps R, S (figure 1) screwed into the ends of the block are coaxial with the duct. Coaxial with two other taps U, V are two holes meeting the duct at right angles. The plane RSUV is horizontal. The duct runs into H, a central vertical pit $^5/_{16}$ inch in diameter. From the centre of the face opposite to UV a horizontal hole J, $\frac{1}{8}$ inch in diameter, is bored into H. The outer end of J may be conical, as in figure 2, to allow free passage to the light. Figure 2 is a vertical section through the axis of J. A window W_1 of plane glass of good quality is set in a recess at the bottom of a larger recess turned out of the upper part of the block. A second window W_2 is set in a recess coaxial with J. The windows are secured with wax. Light from a 6-volt motor headlight lamp M is concentrated by a lens L and, when H contains smoke, a small volume of the smoke is brilliantly illuminated. The distance HL is 1·8 and HM is 4·3 inch. The lamp and lens are mounted on the base which carries the manometer block. A vertical microscope, with objective O, is used for viewing the smoke. The Brownian movement is well seen. The microscope is fixed to a carriage which slides on two vertical rods springing from the base. Its magnifying power is 40; a greater power—say 80—would be better.

---

* The experiments form good exercises for students and are described with that end in view. Unsteady movements of the air near the apparatus and variations of its pressure and temperature bar any claims of the methods to high precision.

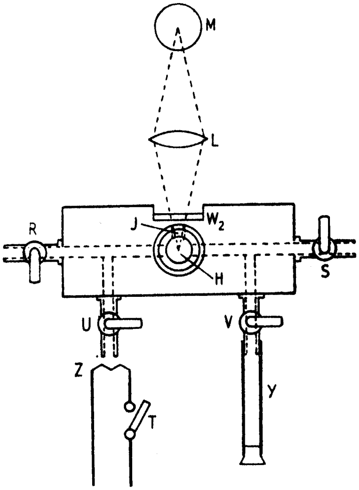
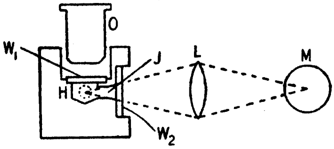


Figure 1.

Figure 2.



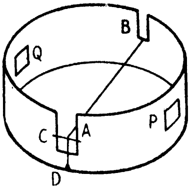


Figure 3.

Figures 1, 2, 3. Smoke manometer.

When smoke is to be drawn into the manometer, taps R, S are closed and U, V are opened. A small cautery Z, which is heated by a current when the switch T is closed, is placed near U, and a match-stalk is held in contact with Z. A rubber tube Y, closed at one end, is fixed to V. If Y be squeezed and then released, smoke is drawn into the manometer duct. Then U, V are closed.

If the pressure in a tube leading to R be greater than in a tube leading to S, the smoke will drift in the direction from R to S, when R and S are opened.

A difference of pressure of 0·1 dyne cm.$^{-2}$ is easily detected. A difference of 2 or 3 dyne cm.$^{-2}$ quickly sweeps the smoke out of the duct, if the taps be fully open.

Messrs. Rotherham and Sons, Ltd., Coventry, supplied their No. 54 taps. These have a bore of $\frac{1}{8}$ inch and have been tested to 75 lbs. inch^{-2} of air pressure. It is essential that the taps and the windows should not leak.

A "cross-wire", by which the drift of the particles may be judged, is stretched across a short length of tube (figure 3) fitting into H. A slot A, corresponding to J, allows passage to the light. Diametrically opposite is a slot B of suitable length. A fine wire C is soldered across A at the height of the axis of J. One end of a silk fibre is secured by wax at D; the fibre passes over C and goes to the bottom of slot B, and the other end is secured with wax as at D. Flats filed on the outer surface of the tube at A and B leave space for the wire C and for the wax. When illuminated, the fibre appears as a bright line. A pair of openings P, Q, diametrically opposite each other, correspond to the duct from R to S.

§ 2. THEORY OF BALANCING COLUMNS

Let ABC, DE be two narrow tubes, bent as in figure 4. Let C, E be at the same level as the horizontal duct of the smoke manometer M, with which the tubes communicate through the taps R, S. The ends A, D are open to the atmosphere. The medium in DE is coal gas of the town supply, at atmospheric absolute temperature θ, that in the non-horizontal part of AB is air heated to θ' by a steam jacket surrounding the effective part of AB, and that in BC is air at temperature θ. The conditions in the horizontal parts of the tubes are governed by the principle that, if the medium be at rest, a *horizontal* line, curved or straight, drawn in the medium and passing, in different parts of its length, through hot or cold air or through gas or a mixture of air and gas, is a line of equal pressure. The air tube is shown with an extra piece BC to meet the case where it is not convenient to have B at the same level as R.

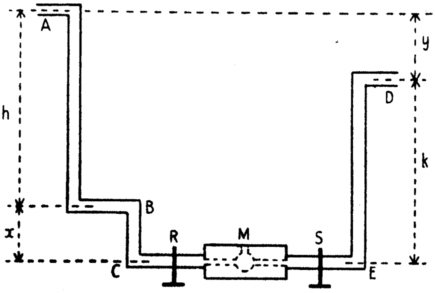


Figure 4. Diagram of balancing columns.

Let $[A]-[B]$, the height of A above B, be h, let $[B]-[C]=x$, $[A]-[D]=y$ and $[D]-[E]=k$.

If ρ and σ be the densities of air and of gas at pressure p and absolute temperature θ, we have $\rho = p/\boldsymbol{R}\theta$, $\sigma = p/\boldsymbol{S}\theta$, where $\boldsymbol{R}$, $\boldsymbol{S}$ are constants. If z be measured downwards, we have, in air at temperature θ,

$$dp = g\rho dz = gp dz/\boldsymbol{R}\theta.$$

Hence $dp/p = gdz/\boldsymbol{R}\theta$, and thus, if θ be independent of z and if p_0 correspond to $z = 0$,

$$\log p - \log p_0 = gz/\boldsymbol{R}\theta. \qquad \ldots\ldots(1)$$

When we apply (1) to the air in AB and BC, we find

$$\log p_B - \log p_A = gh/\boldsymbol{R}\theta', \quad \log p_C - \log p_B = gx/\boldsymbol{R}\theta,$$

and thus

$$\log p_C - \log p_A = g(h/\boldsymbol{R}\theta' + x/\boldsymbol{R}\theta). \qquad \ldots\ldots(2)$$

For the air between A and D and for the gas in DE we have

$$\log p_D - \log p_A = gy/\boldsymbol{R}\theta, \quad \log p_E - \log p_D = gk/\boldsymbol{S}\theta,$$

and thus

$$\log p_E - \log p_A = g(y/\boldsymbol{R}\theta + k/\boldsymbol{S}\theta). \qquad \ldots\ldots(3)$$

When the manometer is balanced, $p_E = p_C$; then, by (2) and (3),

$$k/\boldsymbol{S}\theta + y/\boldsymbol{R}\theta = h/\boldsymbol{R}\theta' + x/\boldsymbol{R}\theta. \qquad \ldots\ldots(4)$$

Now $x + h = y + k$, and thus $x - y = k - h$. Then

$$\frac{\boldsymbol{R}}{\boldsymbol{S}} = \frac{\theta}{k}\left\{\frac{x-y}{\theta} + \frac{h}{\theta'}\right\} = \frac{\theta}{k}\left\{\frac{k}{\theta} - \frac{h}{\theta} + \frac{h}{\theta'}\right\}$$

$$= 1 - \frac{h}{k}\cdot\frac{\theta' - \theta}{\theta'}. \qquad \ldots\ldots(5)$$

This exact result does not involve logarithms.

The constants $\boldsymbol{R}$, $\boldsymbol{S}$ are independent of temperature and pressure. If we specify that ρ and σ are measured at a temperature and a pressure which are the same for both air and gas, we have $\sigma/\rho = \boldsymbol{R}/\boldsymbol{S}$, and then, by (5),

$$\frac{\sigma}{\rho} = 1 - \frac{h}{k}\cdot\frac{\theta' - \theta}{\theta'}. \qquad \ldots\ldots(6)$$

Figure 4 may suggest that the main parts of the tubes are vertical. But they need not be, and are not, vertical; it is only differences of level that count.

§ 3. THE APPARATUS

The general arrangement is shown diagrammatically in figure 5. Free communication at D between coal gas and atmosphere is essential, but, unless it be suitably controlled, the column of gas will be contaminated with air. The outlet from gas to atmosphere is a small hole, centre D, in the end of the tap F. The tap is fixed to one end of the chamber G, a horizontal tube 24·5 cm. in length and 1·7 cm. in diameter. At the other end is a nipple for a rubber tube T_1. By the principle stated in § 2, a horizontal line passing from the atmosphere through D into G is a line of constant pressure. The large volume of G ensures that contamination by air of the gas in T_1 is very slow.

The tube T_1 joins G to the arm S_1 of a three-way piece. The arm S_2 is joined by a short rubber tube to the corresponding nipple of the manometer M. The centre of the hole in this nipple is E. The third arm is connected with Q, the gas supply tube, through the tap S_3.

If the taps S_3, F be open, gas will fill T_1 and G and will stream out at D. In 10 or 15 seconds, all the air will be expelled, and then F is closed. When the manometer test is to be made, tap S_3 is closed and F is opened.

The greatest height of D above E is about 50 cm. If D were vertically above E, the chamber G would be in the way of the observer when he uses the microscope. A plane surface, e.g. a strip of plate glass, provided with screws, is levelled and the heights of D and E above this surface are measured. If these be d and e, then $k=d-e$.

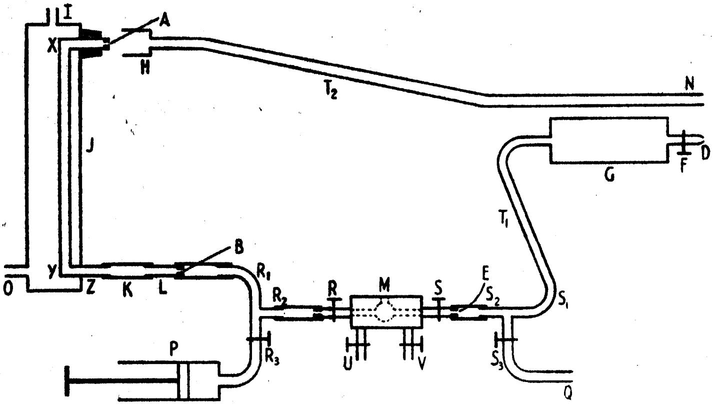


Figure 5. Buoyancy-tube method.

The chamber is supported by a carriage which slides on and can be clamped to a vertical rod. This provides a coarse adjustment of height. A fine adjustment relative to the carriage is effected by a screw.

The greater part of the bent tube AXYZ is surrounded by a steam jacket J, which has efficient lagging.* Steam enters at I and leaves by the tube O, which is coaxial with the part YZ of the inner tube. At the end of XA is a small hole of centre A. The tube YZ is prolonged by a rubber tube K and a metal tube L; both tubes are horizontal. At the end of L is a small hole of centre B. The rubber tube K ensures that L is very nearly at atmospheric temperature. Tube L is connected by a rubber tube to the corresponding nipple of the manometer M through the arms R_1, R_2 of a three-way piece. The third arm, in which is a tap R_3, leads to a small pump P. By P, air can be driven through BYXA to

* A full description is given in paper A cited in § 1.

clear out any coal gas which, through unskilful manipulation, may have penetrated M and entered the tube R_2A. The tap R_3 is closed when the balance is being found.

Tubes O, Z act as trunnions and turn in bearings attached to a length of ⊥-girder. The height of A above B, viz. h, is varied by turning the tube about OZ as axis. The edge of the girder's fin, which has been planed, is levelled. An adjustable strut rising from the girder supports the tube at any desired slope. The heights of A and B above the levelled edge are a and b, and $h = a - b$.

A horizontal line drawn through B and lying within LY is a line of equal pressure, and similarly for AX. Thus, if XA, YB be horizontal, $p_X = p_A$ and $p_Y = p_B$, and the theory of § 2 is applicable.

The temperature of the air near the apparatus is found by a thermometer for each comparison of k with h. The temperature may rise 2 or 3 degrees in half an hour in a class-room. The thermometer should be so placed that it is not heated by radiation from the boiler or by radiation from or by contact with any object, such as the table, which is heated by the boiler. The air itself is little heated by radiation, and slight draughts tend to minimize local rises of air temperature.

§ 4. ANTI-DRAUGHT DEVICE

The apparatus is sensitive to draughts. If, for instance, taps R, S be closed and U, V be open, a to-and-fro movement of a finger near the end of U, or even near the end of a tube a metre in length attached to U, will cause a corresponding movement of the smoke. In a busy class, the movements of students cause draughts which affect the tube A and the tap F unequally, and so lead to spasmodic drifts of the smoke. The difficulty is overcome by a simple device. A short piece of rubber tube fits on the nipple A of the heated tube. One end of a horizontal tube H, of metal or, better, of fibre or other poor conductor, can fit over the rubber; to the other end of H is joined a rubber tube T_2 which leads to an open nozzle N near the tap F of G. The axis of N is parallel to that of F and the ends of N and F are not more than 2 or 3 cm. apart. The part of T_2 near H is horizontal for a little distance, and care is taken that the remainder of T_2 is at atmospheric temperature. Then, when H is in position on A, the pressure at A inside H equals that of the atmosphere at the same level, unless there be large streaming motions of the air in the room. The device is so effective that violent local aerial disturbances near N and F produce only minute and transient movements of the smoke.

§ 5. TEMPERATURE OF STEAM

The temperature T° C. of the steam in the jacket can be found from a steam table. This gives T as a function of the pressure as measured by a barometer, with its mercury at 0° C., at sea level in latitude 45°, where $g_N = 980{\cdot}62$ cm. sec.$^{-2}$ At the Cavendish Laboratory, $g_L = 981{\cdot}27 = 1{\cdot}000663 g_N$. If μ_1 and μ_0 be the density of mercury at t_1° C. and at 0° C., $\mu_1 = \mu_0(1 + 0{\cdot}000182 t_1)^{-1}$. If a pressure, which gives height H_N cm. in latitude 45° with mercury at 0° c., give H cm. at the Laboratory with mercury at t_1°c. we have $g_N\mu_0 H_N = g_L\mu_1 H$, and thus

$$H_N = H(1+0{\cdot}000663)(1+0{\cdot}000182t_1)^{-1} = H + H(0{\cdot}000663 - 0{\cdot}000182t_1).$$

In the small term, we put $H=76$. Then

$$H_N = H + 0{\cdot}050 - 0{\cdot}014t_1.$$

For H_N near 76 cm., the boiling point rises by 0·37° c. when H_N rises by 1 cm., and $T=100+(H_N-76)\times 0{\cdot}37$. Thus

$$T = 100 + (H - 75{\cdot}95)\times 0{\cdot}37 - 0{\cdot}0052t_1. \qquad \ldots\ldots(7)$$

If the temperature of the air near the apparatus be t° c., the absolute temperatures of the air and of the steam are $\theta = t + 273{\cdot}2$, $\theta' = T + 273{\cdot}2$.

§ 6. EXPERIMENTAL DETAILS

The boiler, containing water, is started. The girder and the glass plate are levelled and b and e are measured. The height a of the upper nipple of the hot-air tube above the girder is found, and then the anti-draught device is adjusted. When steam is passing freely, the cell is charged with smoke and taps U, V are closed. Gas is allowed to flow through S_3 and F, and thus the chamber G is filled with gas. Then S_3 and F are closed. When a test for balance is to be made, F is opened. Then, R being open, S is opened cautiously; the direction of drift shows whether G should be raised or lowered. When the balance has been found, the height d of D above the plate is measured. If, when one pair of values of a–b or h and d–e or k has been found, k be set against h on a slide rule, approximate values of k can be read off for other values of h, and thus time can be saved in reaching the desired balances. The values of h should be large so as to minimize the effects of zero errors.

For each balance of k against h, the air temperature is observed. The boiling point T° must also be known, but, as T rises by no more than 0·1° for a rise (0·27 cm.) of H_N from 76 to 76·27 cm. or for a fall of 19° in t_1, the temperature of the mercury, the variations of T during the experiment are usually negligible.

A test was made of the efficiency of the chamber G (figure 5) in preventing contamination of the gas column by air. The effective height h of the hot-air column was 83·45 cm. The tube T_1 and the chamber G were filled with gas, and then the values of k needed for a balance were found at various times after the filling. The tap F was kept open throughout. The changes with time of the apparent value of σ/ρ are shown below.

Time (min.)	0	1	3	5	7	9	13	17
10 σ/ρ	5·02	5·02	5·04	5·06	5·10	5·11	5·21	5·32

The results show that no error due to contamination with air need be feared if the balance be found within two minutes of filling the chamber with gas. The test was made a few days before the measurements recorded in § 7. The ratio σ/ρ varies from day to day. At the Gas Works, σ/ρ is found by the method of effusion through a very small orifice.

§ 7. PRACTICAL EXAMPLE

Miss Mary Stead, of Girton College, made the observations. Conditions were favourable. The temperature of the air near the apparatus was steady, and no disturbance due to draughts was observed; the anti-draught device was not used. The barometric height was initially 76·91 cm.; the effect on the boiling point of a slight fall in H was negligible. The temperature of the barometer was 13·9° c. By (7),

$$\theta' = 273{\cdot}2 + 100 + (H - 75{\cdot}95) \times 0{\cdot}37 - 0{\cdot}0052 \times 13{\cdot}9 = 373{\cdot}5^\circ \text{ K.}$$

For each h, four independent settings and readings for k were made. The mean divergence, without regard to sign, from the mean k, of the four values of k found for a given h is denoted by Δ. The values of $1000\Delta/k$ in the table are small.

h	k (mean)	$\dfrac{1000\Delta}{k}$	θ	$\dfrac{\theta'-\theta}{\theta'}$	$\dfrac{h}{k}$	$\dfrac{\sigma}{\rho}$
(cm.)	(cm.)		(°K.)			
89·70	41·300	0·61	287·2	0·23106	2·1719	0·4982
87·00	39·825	0·63	287·3	0·23079	2·1846	0·4958
81·15	37·262	1·01	287·3	0·23079	2·1778	0·4974
77·05	35·338	1·24	287·3	0·23079	2·1804	0·4968
72·60	33·412	1·12	287·3	0·23079	2·1729	0·4985

The mean value of σ/ρ is 0·49734, with a "probable error" of 0·00033.

§ 8. WHIRLING-ARM METHOD

In place of the column of hot air, the whirling arm described in §§ 3, 4 of Paper A may be used to give a defect of pressure. In figure 6, BC is the arm, which turns about the vertical axis OZ.

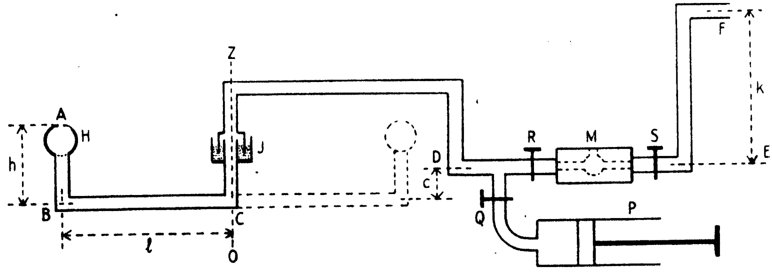


Figure 6. Whirling-arm method.

The distance of the test hole A in the static-head tube H from OZ is l, and its distance above BC is h. There is a liquid joint J in CD. The line DE through the smoke manometer M is horizontal, and BC is at distance c below DE. The tube EF contains coal gas; the distance of F above DE is k. The details of the arrangement on the gas side are shown in figure 5. The pump P of figure 5 is connected, through the tap Q, to the tube from D to M, and serves to expel any trace of coal gas from AD.

If the absolute temperature of the apparatus and of the surrounding air be θ, we have, by § 2,

$$\log p_B - \log p_A = gh/R\theta, \qquad \log p_C - \log p_D = gc/R\theta. \qquad \ldots\ldots(8), (9)$$

Let ω be the angular velocity of BC about OZ. Then in BC

$$dp/p = \rho r\omega^2 dr/p = r\omega^2 dr/R\theta,$$

and thus
$$\log p_B - \log p_C = \tfrac{1}{2}l^2\omega^2/R\theta = \tfrac{1}{2}v^2/R\theta, \qquad \ldots\ldots(10)$$

where v is the velocity of the test hole A along its path. From (8), (9), (10),

$$\log p_D - \log p_A = [g(h-c) - \tfrac{1}{2}v^2]/R\theta. \qquad \ldots\ldots(11)$$

If the static-head tube H be suitably corrected (Paper A, § 4), the pressure at A inside H equals that in the atmosphere at a distance from A and at the same level as A.

On the other side of the manometer, we have, for the air and the gas respectively,

$$\log p_A - \log p_F = g(k+c-h)/R\theta, \quad \log p_E - \log p_F = gk/S\theta, \quad \ldots\ldots(12), (13)$$

and thus
$$\log p_E - \log p_A = g(h-c-k)/R\theta + gk/S\theta. \qquad \ldots\ldots(14)$$

When the manometer shows a balance, $p_E = p_D$. When we subtract (14) from (11), and then multiply by θ, we have

$$gk/R - gk/S - \tfrac{1}{2}v^2/R = 0,$$

and thus
$$\frac{R}{S} = 1 - \frac{v^2}{2gk}. \qquad \ldots\ldots(15)$$

This equation takes the place of (5) in § 2. If ρ, σ be measured at a temperature and a pressure which are the same for both air and gas, $\sigma/\rho = R/S$, and thus

$$\frac{\sigma}{\rho} = 1 - \frac{v^2}{2gk}. \qquad \ldots\ldots(16)$$

No knowledge of the temperature is needed in (16), but uniformity of temperature is assumed.

The use of the whirling arm involves the correction of the static-head tube required for the small velocities employed, and this correction has to be made by running the whirling arm against a heated buoyancy tube, as described in § 11 of Paper A. The anti-draught device (§ 4) cannot be employed with the whirling arm. For these reasons, it seemed best in the present paper to treat the method of the buoyancy tube as standard rather than that of the whirling arm.

§ 9. DETERMINATION OF σ/ρ BY WEIGHING

The results found by the previous methods may be tested by weighing. A vessel of mass M and internal volume V is weighed when it is filled with (1) air and (2) coal gas at the temperature and pressure of the surrounding atmosphere. In (1) the mass of the vessel and its contents is $M + \rho V$, and in (2) is $M + \sigma V$. The balancing load may consist of brass weights and aluminium weights. We suppose that the brass weights are the same in each case, and that the balancing load for (1) exceeds that for (2) by D gm. of aluminium of density α. The effective mass of D gm. of aluminium is $D(1-\rho/\alpha)$, where ρ is the density of the air in the balance case. The buoyancy effects of the air on the vessel and on the brass weights are the same in both cases. Hence

$$(\rho-\sigma)V=D(1-\rho/\alpha),$$

and thus

$$\frac{\sigma}{\rho}=1-\frac{D}{V}\left(\frac{1}{\rho}-\frac{1}{\alpha}\right). \qquad \ldots\ldots(17)$$

An aluminium hot-water bottle B (figure 7), of about 1500 c.c., may be used. To its mouth is fixed, by marine glue or other adhesive, a plate C through which pass, with soldered joints, two tubes L and S. Tube L goes nearly to the bottom of B, and S only just passes through C. The tubes project a little distance above C. Corks are provided for them. The temperature of B is assumed to be that of the atmosphere, and hence it should not be touched by the hand. It may be lifted by a rod inserted into a loop H soldered to C. Care must be taken that the bottle does not pick up dirt of any kind.

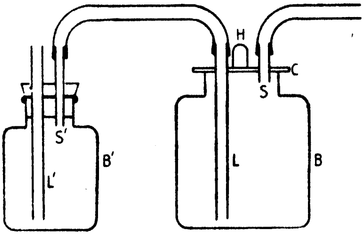


Figure 7. Weighing method.

For the air filling, a tube from the blowpipe bellows (not from the observer's mouth—Oh, dear, what things they do!) is fixed to L, and the bellows are worked until it is certain that any gas remaining after a previous gas filling has been expelled. Then the bottle, with the corks, is put into the balance case; the tubes L, S are left open. After a few minutes to allow the bottle and its contents to attain the temperature of the air in the case, the weight is found—best by the method of oscillations.

The bottle is now filled with gas. A tube leading from the gas supply is attached to S, and gas is allowed to flow for a little time to expel all air from B. If the gas come from a main near heating pipes, its temperature may be above that of the room. If L and S be corked immediately after the tube is removed from S, the pressure of the contained gas will be below that of the atmosphere when its temperature has sunk to that of the atmosphere, or, if the corks leak, air will flow in as the gas cools. To meet this difficulty, a bottle B′, with tubes L′ and S′, is put in tandem with B, as in figure 7; tube S′ is connected to L. After gas has streamed through B and B′ for a sufficient time, the tube from the supply pipe is removed from S, which is immediately corked. The pair of bottles is then moved to a place near the balance case. If the gas cool, air will flow into B′ through L′, but the small volume of gas which flows into B from B′ will be only slightly contaminated with air. After a few minutes, the tube connecting S′ to L is removed, and L is immediately corked. Then the bottle B, with the contained gas, is weighed.

Before the plate C is fixed to B, the volume V_0 of B to the level of the top of the neck is found by weighing it (1) empty and (2) filled with water at a known temperature. Allowances are made for the buoyancy of the enclosed water and of the balancing masses. The coefficient of cubical expansion of aluminium ($7{\cdot}65 \times 10^{-5}$) is so small that changes of volume due to moderate changes of temperature are negligible. The brass tubes L and S are weighed before they are soldered into C, and the volume V_1 of brass below C is estimated. The small volume of the tubes between the corks and the level of the lower face of C is V_2. Then $V = V_0 - V_1 + V_2$.

REFERENCE

SEARLE, G. F. C., 1941. "A whirling arm for producing small differences of pressure, its calibration by aid of a smoke manometer and its application to measurements of surface tension." *Proc. Phys. Soc.* **53**, 681.

DEMONSTRATION

PHOTOMICROGRAPHS illustrating the paper "An achromatic reflection microscope for use with visible or ultra-violet light" (*Proc. Phys. Soc.* **53**, 714 (1941)). *Demonstrated* 6 *February* 1942 *by* B. K. JOHNSON.

IN the above paper, the writer described an achromatic reflection microscope and gave details of the computed aberrations of the optical system. Photographs taken with the completed instrument were submitted with the original paper, but were not published at the time; an opportunity was, however, given to show them at a later meeting, and (because of the interest which they aroused among those Fellows able to attend) the author was subsequently invited to prepare a short description of the demonstration, illustrated by the photomicrographs in question. This description follows, though the original text should be consulted in order to appreciate the full meaning of each of the photomicrographs.

Figures 1 and 2 illustrate the definition obtainable with the reflection microscope designed for use with monochromatic light (λ 5461 A.) employing glass for the optical system (see *J. Sci. Instrum.* **11**, 384 (1934)). The instrument was originally intended for the observation of opaque objects in order to overcome back-reflection defects, but it may be used for transparent objects as shown in these two photographs.

Figures 3 (*a*), (*b*) and (*c*) show an opaque object (CuS) photographed with the achromatized reflection microscope, but using light of wave-lengths 5461 A.,

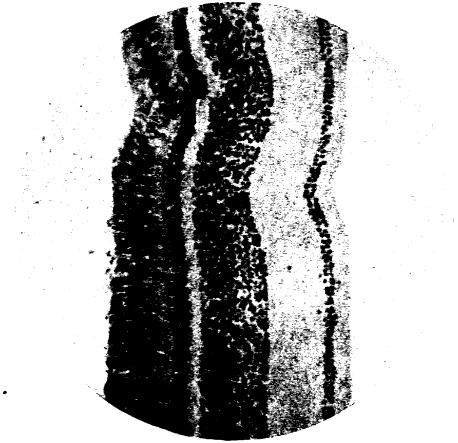


Fig. 1. Retina of owl.
λ=5461A. N.A.=0·31.
M=105x.

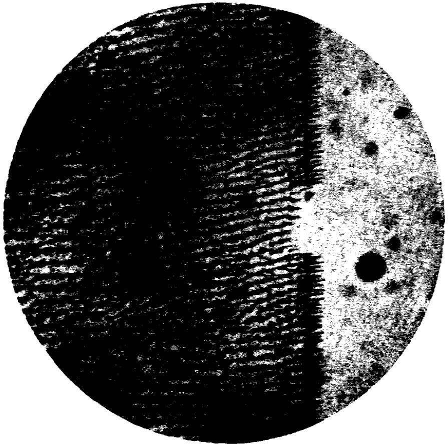
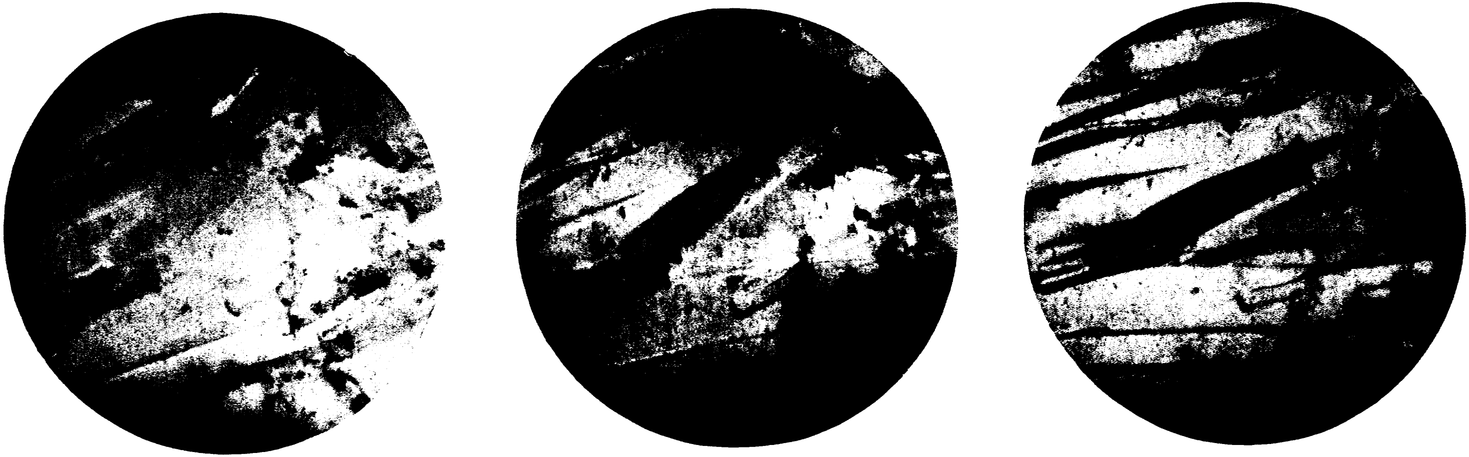


Fig. 2. Wing of insect.
λ=5461A. N.A.=0·31.
M=105x.

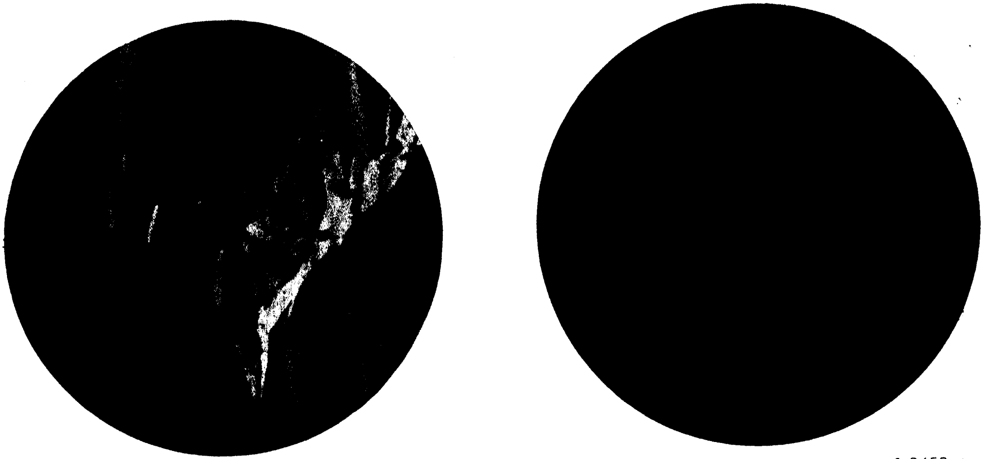


(a) λ=5461A.

(b) λ=3651A.

(c) λ=2749A.

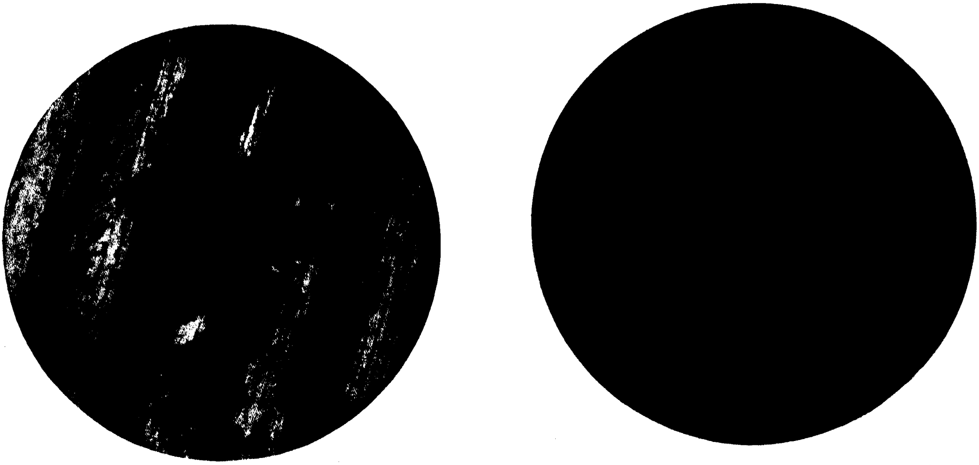
Fig. 3. Differential reflection effects when using various U.V. wavelengths. Object: CuS. N.A.=0·47. M=140x.



Specimen (zinc blend, Tsumed, S.W. Africa), illuminated and photographed with white light.
M = 140x. N.A. = 0·47.

Same specimen, illuminated by light of 3650 to 3131A and photographed in colour by the fluorescent light emitted.
M = 140x. N.A. = 0·47.

Fig. 4.



Specimen (an aluminium alloy) illuminated and photographed with white light.
M = 140x. N.A. = 0·47.

Same specimen, illuminated by light of 3650A to 3131A and photographed in colour by the fluorescent light emitted.
M = 140x. N.A. = 0·47.
N.B. A filter (transmitting no light below 4000A) was mounted in the plane of the exit-pupil of the eyepiece. Exposure = 75 hours

Fig. 5.

Figs. 4 and 5. Comparison photomicrographs taken with B.K.J. reflection microscope showing benefit of fluorescent effects made available by the optical system of the instrument.

3651 A. and 2749 A., respectively. In this illustration there are two points worthy of note: firstly, no alteration in the fine adjustment of the microscope has been made for the three photomicrographs, which it will be seen are in equally good focus. Secondly, the effects of differential reflection (when using light of various ultra-violet wave-lengths) can be clearly seen.

Figures 4 and 5 show the possibilities of improving contrast and of detecting impurities by utilizing fluorescent effects.

In each pair of photomicrographs the monochrome picture gives the effect seen when white light only is used, whilst the coloured illustration shows the appearance of precisely the same part of each object when it is illuminated by ultra-violet light of wave-length 3650 A. to 3131 A. and seen by the fluorescent light emitted by the specimen. These photomicrographs in colour are only intended to convey what the eye actually sees when looking into the microscope, and it is not suggested that such photographs should be taken for all specimens which fluoresce, for the exposure required in this case (namely 75 hours) would in general be prohibitive. They do serve to show, however, that here is an additional means of aiding the possible identification of the components of an alloy (for example), or in the detection of impurities; and that the instrument described in the original paper is well suited to this purpose.

DISCUSSION

On papers by W. E. BENHAM (*Proc. Phys. Soc.* **54**, 121 (1942)) and G. D. YARNOLD (*Proc. Phys. Soc.* **54**, 47 (1942)) entitled " On the nature of temperature" and "Notes on electric and magnetic dimensions", respectively.

Professor DINGLE. I think Mr. Benham is in error in assuming that dimensions are associated with the ***nature*** of temperature: they characterize the ***method of measuring*** temperature, and this is arbitrary. Thus, if we define temperature as energy per unit mass, temperature will have finite dimensions in length and time, because a change in the units of measurement of length and time will cause a change in the measured value of a temperature. On the other hand, if we define temperature as a thermometer reading, the degree being 1/100 of the relative expansion of a standard substance between the temperatures of melting ice and boiling water, then the measured value of a temperature is unaltered by a change in any other unit, so that temperature has an independent dimension of its own. There is no meaning in asking what are the " true " dimensions until we have decided which is the " true " method of measuring temperature.

Much confusion is caused by the tacit assumption that the two methods are equivalent. An outstanding instance of their divergence is found in the temperature of interstellar space. Expressed as energy per unit mass, the temperature is some thousands of degrees, but a thermometer placed there would read about 3° absolute. The molecules move very fast, but the matter is so rarefied that the energy which the thermometer receives by impact is radiated as fast as it is received.

There is a similar confusion in the measurement of time. If the unit of time is the mean solar second, or sidereal day, the measurement of time is unaffected by the measurement of anything else, and time has dimensions of its own. The Earth then rotates uniformly by definition. Astronomers, however, now believe that the Earth is "slowing down". This means that they must tacitly adopt another standard of time, which on investigation proves to be the standard according to which Newton's first law of motion (with light substituted for an undisturbed body) is automatically true. On this standard equal times are those in which light moves over equal lengths, and a change in the measurement of length therefore affects the measurement of time; hence time has finite dimensions in length.

Much of the confusion in this subject arises from oversight of the fact that dimensions characterize the arbitrary method of measuring physical quantities, and not the "ultimate nature" of the quantities.

Dr. E. A. GUGGENHEIM. The object of the following remarks is to indicate how a little extra care in our choice of terminology could bring to an end the hitherto ceaseless confusion between quantities with physical dimensions on the one hand and pure numbers on the other. A quantity with physical dimensions is one whose numerical measure depends on the units used, whereas a pure number is obviously independent of these units. This statement can be illustrated by two familiar examples.

If we define density as mass per unit volume, then the density of water is (in round figures) 1 kg./l. or 62½ lb./ft.^{3} Again, if we define heat capacity as the energy required (under certain specified conditions) to produce unit rise of temperature in unit quantity of a substance, then the heat capacity of water is (in round figures) 1 B.Th.U./lb. deg. F., or 4·2 joules/g. deg. C., or 18 cal./mole deg. C. If, further, we define specific gravity as the ratio of the density of a substance to that of water, then the specific gravity of mercury is 13·6. Again, if we define specific heat as the ratio of the heat capacity of a substance to that of water, then the specific heat of mercury is 0·033. According to the chosen definitions it is evident that density and heat capacity have physical dimensions, while specific gravity and specific heat are pure numbers.

Quantities with physical dimensions, as well as pure numbers, may conveniently be denoted by symbols. For example, if $\mathbf{c}$ denotes the speed of light in empty space, then (in round numbers) $\mathbf{c} = 3 \times 10^8$ m./sec. = 186,000 miles/sec. = 1 light-year/year, and for $\mathbf{c}$ so defined it is wrong to write $\mathbf{c} = 3 \times 10^8$ (or any other number).

Turning now to electromagnetic quantities, if f denotes the force of repulsion between two electric charges q_A and q_B in free space at a distance apart r_{AB}, then according to Coulomb's law the ratio $q_A q_B / f r_{AB}^2$ is a universal constant, which we can denote by $\boldsymbol{a}$. A short time ago I suggested (Guggenheim, 1941)* that the charge which repels an equal charge at a distance of 1 cm. in empty space with a force of 1 dyne be called a *franklin*, and I shall use this terminology. We have, then, for the universal constant $\boldsymbol{a}$ (in round numbers),

$$\boldsymbol{a} = 1\,\frac{\text{franklin}^2}{\text{erg cm.}} = \frac{1}{9\times 10^9}\,\frac{\text{coulomb}^2}{\text{joule m.}} = \frac{1}{9\times 10^9}\,\frac{\text{sec.}}{\text{ohm m.}},$$

and it is evident that $\boldsymbol{a}$ has physical dimensions.

If we denote current elements by $i\,ds$, magnetic poles by p and the velocity of light by $\boldsymbol{c}$, all these quantities being measured in any units whatever, then we can construct table 1 for forces of attraction in free space and in a medium. The quantities ϵ and μ are characteristic constants of the medium and, since by the chosen definition each is a measure of the ratio of a force in empty space to a force inside a medium, it is evident that each is a pure number.

Table 1. Forces of attraction between electric charges, current elements and magnetic poles

	Electric charges	Current elements	Magnetic poles
Empty space	$-\frac{q_A q_B}{\boldsymbol{a} r_{AB}^2}$	$\frac{i_A i_B ds_A ds_B}{\boldsymbol{a}\boldsymbol{c}^2 r_{AB}^2}$	$-\frac{\boldsymbol{a}\boldsymbol{c}^2 p_A p_B}{r_{AB}^2}$
Medium	$-\frac{q_A q_B}{\epsilon \boldsymbol{a} r_{AB}^2}$	$\frac{\mu i_A i_B ds_A ds_B}{\boldsymbol{a}\boldsymbol{c}^2 r_{AB}^2}$	$-\frac{\boldsymbol{a}\boldsymbol{c}^2 p_A p_B}{\mu r_{AB}^2}$

In order to avoid ambiguity between the quantities $\epsilon\boldsymbol{a}$ and $\mu/\boldsymbol{a}\boldsymbol{c}^2$, both having physical dimensions, on the one hand, and the pure numbers ϵ and μ on the other, I suggest the adoption of the terminology shown in table 2. In choosing

Table 2. Proposed terminology

$\epsilon\boldsymbol{a}$ Permittivity of medium	$\frac{\mu}{\boldsymbol{a}\boldsymbol{c}^2}$ Permeability of medium
ϵ Dielectric constant of medium	μ Magnetic constant of medium
$\boldsymbol{a}$ Permittivity of empty space	$\frac{1}{\boldsymbol{a}\boldsymbol{c}^2}$ Permeability of empty space
$=1\,\frac{\text{franklin}^2}{\text{erg cm.}} = \frac{1}{9\times 10^9}\,\frac{\text{sec.}}{\text{ohm m.}}$	$=\frac{1}{9\times 10^{20}}\,\frac{\text{g. cm.}}{\text{franklin}^2} = \frac{1}{10^7}\,\frac{\text{ohm sec.}}{\text{m.}}$

---

* References are on p. 296.

this terminology I have been guided by a desire for the maximum precision consistent with the minimum departure from current practice. It will be noticed that the only essentially new terms are the *franklin* and the *magnetic constant*.

Attention has been drawn to the physical dimensionality of the universal constant $\boldsymbol{a}$ in a very instructive paper by Sommerfeld (1935), unfortunately too little known. Sommerfeld used the symbol ϵ_0 where I have used $\boldsymbol{a}$ and ϵ where I have used $\epsilon\boldsymbol{a}$. I think that the notation which I propose has two advantages over that of Sommerfeld's paper. Firstly, I propose that a single symbol $\boldsymbol{a}$ be used for a universal constant: secondly, I propose that the quantities usually listed in tables be also denoted by the single letters ϵ and μ. The greatest change entailed by my proposals is that tables of μ should be renamed tables of *magnetic constants*.

In conclusion, I want to emphasize that, whereas the question whether an unambiguously defined quantity x has physical dimensions or is a pure number has only one answer; when this answer is " Yes ", the further question " What are the dimensions of x ? " has an infinite number of answers, depending on one's personal preference. I have accordingly completely refrained from futile discussion of the latter question.

Dr. G. B. **Brown**. Dr. Yarnold stresses the fact that a constant a must always be included in the equation expressing an inverse-square law of force,

$$F = a\frac{q_1 q_2}{r^2},$$

and says that all we can say about the *dimensions* of a and q is that aq^2 has the dimensions of force multiplied by the square of a distance. Equally, then, according to this argument, all we can say about the *magnitudes* of a and q is that aq^2 will equal the product of the force measured and the square of the measured distance. Thus we never know the magnitude of an electric charge.

It does not seem generally realized that when we get over the second *impasse* by *defining the magnitude* of unit q in such a way as to make a unity, we also get over the first *impasse* by *defining the dimensions* of q in such a way that a is dimensionless. As I have said, " It can then, *and only then*, be dropped from the equation " (Brown, 1941, p. 424).

With regard to the derivation of the form of the inverse-square law in polarizable media: measuring the fall of potential across a condenser when a dielectric is introduced cannot tell us more than the *average* force which may be supposed to act on a unit charge as it is taken from one plate to the other. We certainly do not know the force on it at *any point*. If it is near a molecule we do not know what it is, and if it is not near a molecule, it must be in a cavity of some sort, and we want to know its shape. Then again, when the polarization of the medium is due entirely to two point charges we must remember that it will be differently polarized according as they are like or unlike. I do not think there is any escape from some sort of cavity treatment such as I have attempted in an earlier paper (Brown, 1940).

Prof. R. O. KAPP. I think that in discussions on units and dimensions there is a tendency to confuse three different questions. The first of these concerns standards for purposes of copying, and is the only one to which experimental work has any relevance. The second involves metaphysical considerations. I doubt if it ever has any relevance. The third is concerned with labour-saving. I think this consideration alone ought to govern our choice of unit quantity in physics.

If we were untrammelled by the past we should seek to have only one unit for each physical quantity, and multiply this by sundry powers of ten, when it was inconveniently large or small. But custom and tradition are so strong that we cannot now dispense altogether with a multiplicity of units for certain quantities. We need, for instance, several units of weight and several units of length. The horse-power can hardly be suppressed in favour of the watt, or the calorie in favour of the erg. But our objection to such multiplicity, as well as our arguments in its favour, have nothing whatever to do either with experiment or metaphysics. The objection is that conversion factors are troublesome and lead to mistakes in arithmetic. The argument in favour is that it is less trouble to use conversion factors than to abandon established habits of thought.

This argument is weak when applied to existing systems of units in electricity and magnetism. Here we are saddled with three units for every quantity. We have the abvolt, statvolt, and volt. We have the abfarad, statfarad and farad, and so on throughout the whole of electricity and magnetism. Two of these three systems have been justified on the grounds that they are " natural ". The reason is that units in them are derived from the properties of a vacuum. But why should the properties of a vacuum be called more natural than those of any other medium? Is it merely because of the plentifulness of a vacuum in the Universe?

The real reason for our three units for every quantity has nothing to do with nature. It is because in electricity and magnetism, five unit quantities have been defined arbitrarily, and five is one too many. It is strange that this over-determination is defended by those who argue that the proper number of arbitrary determinations is three. The five which have been made in electricity and magnetism are the units of length, mass, time, permeability and permittivity.

This is the reason why the International Electro-technical Commission recommended unanimously in 1935 that the M.K.S. system should be adopted, in which only four arbitrary units are determined. These are the metre, the kilogram, the second and 10^7 times the permeability of a vacuum. In this system those magneto-electric units which are used most often, such as the ampere, volt, ohm, and farad, are identical with those in the practical system.

In conclusion I would like to point out that the fact that four arbitrary unit quantities have to be determined in electricity and magnetism is sufficient answer to those who argue that all physical quantities can be adequately expressed in terms of less than four fundamental dimensions.

Dr. C. **Fox**. A serious drawback to the theory of dimensions as practised at present is that it fails to distinguish between scalar and vector quantities, and I do not think that the theory can be clarified until modifications are introduced which make this distinction clear. Displacement is a vector quantity, and some vectors such as force, velocity, etc., show their vectorial nature in the theory of dimensions by having unity as their dimension in length. According to the formulae of Mr. Benham, the introduction of " Newton's constant " N gives force a dimension of -2 in length, which completely hides its vectorial nature. I consider this an objection to the theory proposed by Mr. Benham unless some other method of showing the vectorial nature of force is introduced in compensation.

Authors' replies:

Dr. G. D. **Yarnold**. There seems to be considerable confusion between the definition of a physical quantity and the statement of the way in which it is to be measured. Where a concept is so fundamental as to be incapable of definition, it is only possible to state the method to be adopted in its measurement. Thus, the concept *length* is indefinable, but length is to be measured by the ratio of two lengths, one of them a standard. However, this does not make length dimensionless. Dielectric constant, on the other hand, is defined as the ratio of two capacities, and is therefore defined to be dimensionless, but in this case the definition also indicates the method of measurement.

I entirely agree with Dr. Brown's remark about the magnitudes of a and q. But surely we are not at liberty to define the dimensions of q in such a way as to make a dimensionless? Electric charge appears to be something absolutely fundamental in the nature of things, whose dimensions cannot be fixed by reference to length and time. As regards Dr. Brown's second point, I can only reply that I have treated the medium from the classical point of view as continuous. At this stage there is no reason to enquire why a collection of atoms behaves as a medium.

I have read Dr. Guggenheim's comments and feel that his suggestions are thoroughly sound.

In reply to Professor Kapp, I agree that the definition of four, but not more than four, fundamental quantities is necessary in electromagnetism, but I regard the choice of the permeability of a vacuum in the M.K.S. system as regrettable.

Mr. W. E. **Benham**. I am interested in Dr. Fox's point of view, and agree that the theory of dimensions would be improved if it were possible thereby to distinguish between scalars and vectors. Dr. Fox fails to establish his objection to the point of view of our paper, since he provides no criterion whereby the vectorial nature of a vector shall be judged. He starts by saying that the ordinary theory is deficient in the very respect for which he cites objection to the modification proposed. For his objection to hold it would be necessary that not merely *some* vectors . . . . have their dimensions in length unity, or, shall we say, odd, but *all* vectors. Conversely, scalars must presumably have *even* dimensions in length. Examples of vectors having even and scalars having odd length dimensions may readily be found. Thus a vector area $\boldsymbol{S}$ has dimensions L^2; similarly, regarding zero as an even number, the angular velocity ω, of dimensions $1/T$, regarded as a vector drawn in right-handed manner along the axis of rotation, is of even dimensions in L; while energy per unit volume (a scalar, even when numerically equal to the vector pressure) has odd dimensions in L.

Now on our theory force has dimensions M/L^2. There would appear to be a case for including a unit vector $\boldsymbol{r}_1$, parallel to the direction of the force. More succinctly, if we regard the force as acting normally on a small area S,

$$[F] = \frac{M}{\mathbf{S}},$$

where $\boldsymbol{S}$ is the vector area. Some means of telling the direction of $\boldsymbol{F}$ would thus seem to be provided by regarding force as *mass per vector area*. This seems to be a real advance, for what could more aptly describe our *sensation* of force than this? I must express my acknowledgements to Dr. Fox for putting forward his point about vectors, without which the above dimensional expression for force might have been passed over.

In reply to Professor Dingle, it will be evident that we share Lanchester's views (1936) that "in physics the term dimensions denotes the kind or quality of a physical entity" and "that a single physical entity may possess more than one dimensional value is to the author unthinkable". While it is hardly within the scope of physics to discuss the nature of the fundamental units, physics can describe quantities not regarded as fundamental, such as temperature, in terms of the fundamental units. To this extent it is legitimate to talk of the nature of temperature. The fact that on the basis of thermometry temperature has dimensions of its own should make us suspect that thermometry does not go to the root of the matter. We then turn to the kinetic theory and to radiation, with the results indicated in the paper. It might, of course, be contended that one might choose to regard temperature as more fundamental than, say, either length or time. It would then be an error to talk of the nature of temperature, but one might perhaps talk of the nature of time if it were possible, on this basis, to reduce time to temperature measure. Prof. Dingle's suggestion for the measurement of temperature in interstellar space seems to be similar to the idea of using a sledge-hammer to knock in a tin-tack. It is possible to use the wrong tool. That a thermometer is suitable in laboratory experiments is due to the fact that it is possible to bring the instrument into temperature equilibrium with its surroundings.

Regarding the problem of the variability of the earth's rotation, the standard of time need not be based, as Prof. Dingle suggests, on the distance travelled by light, despite the jeopardy into which the mean solar second is thrown by the supposed variation of the diurnal motion of the earth. For the temperature-controlled master-clocks used at observatories are capable of such precision that it is often the clock which supervises the astronomer rather than the reverse. One might therefore rest content, for example, with the definition of sidereal time as determined by star transit methods in any particular year (say 1942 or 2000 A.D.) and conserved in precision clocks until such time as the amazing refinements now attainable should perforce be superseded. Any slowing down of the earth's rotation (its amount is of the order of ½ second per century) would then show up as a gradual divergence between astronomical and horological readings. Even if the ability of clocks to maintain regularity over long periods be in question, there are devices for overcoming this, and in view of compensations we should be adhering to a method of measurement which does not involve change of the time unit with length. Indeed, since length and time are separate fundamental units, measurement of time cannot involve length, though it may involve the ratio of two lengths or the coincidence of two marks. For example, the fact that we certainly use a length to measure temperature in a mercurial thermometer has no connection with the reciprocal length which, I am suggesting, connotes the intrinsic significance of temperature—the "thinness" of a pulse of electromagnetic radiation.

With regard to Professor Kapp's communication, as this is mainly directed to Dr. Yarnold, my comments will be but brief. From the point of view of the ordinary practising physicist, the question of labour-saving is of paramount importance, and choice of units should, indeed, be guided, if not governed, by this consideration. We are in favour of the M.K.S. system, and would also suggest rational units, making the so-called "permeability of a vacuum" equal to $(1{\cdot}256)10^{-6}$ henry per metre. There seems to us every reason for at length supporting Heaviside in his elimination of 4π from places where it ought never to have occurred.* The use of M.K.S. units, whether rational or irrational, should not be allowed, however, to influence us into thinking that the attribution of dimensions to μ and κ is anything but an artifice designed to enable us to forget about the presence of force constants.

REFERENCES

Brown, G. B., 1940. *Proc. Phys. Soc.* **52**, 577 ; 1941. *Ibid.* **53**, 418.
Guggenheim, E. A., 1941. *Nature, Lond.*, **148**, 751.
Lanchester, F. W., 1936. *The Theory of Dimensions and its Application for Engineers* (London), pp. vii and 125.
Sommerfeld, A., 1935. *Z. tech. Phys.* **16**, 420.

* Compare Vigoureux, P., *Proc. Phys. Soc.* **52**, 611 (1940).

OBITUARY NOTICES

JULIUS RHEINBERG

Rheinberg was a true amateur in the full sense of the word, for while engaged in ordinary commerce as his means of livelihood he devoted himself in all his spare time to the pursuit of science. His earlier work was with the microscope, and he published several papers in connection with the microscopic image. In particular, he devised a very fascinating way of exhibiting fine structure by means of a diaphragm placed in the lower focal plane of the condenser, of which the central area was one colour and the outer zone a contrasting colour. Thus the more oblique rays forming the image were of a different colour from the axial ones; whilst this probably has no advantage as regards the resolving power of the objective, it often makes the structure easier to observe. At the time when so much discussion took place upon the merits of the Abbe theory of the formation of the microscopic image, Rheinberg made a large number of experiments on diffraction in the endeavour to come to a positive conclusion on the points involved.

In the last war, he worked out a method of constructing fine graticules by photography which was of great value to the country—a method which the Government requested him to keep and work as a secret process. He once told me that he had obtained the fundamental idea from some of the early and forgotten methods of producing a photographic image—such as those, for example, described in Hunt's *Photography* (1854). These graticules were a great improvement on anything that had been made up to that time.

More recently he had turned his attention to stereoscopy. He made a number of experiments, but has published nothing of his results.

He was a regular supporter both of the Royal Microscopical Society and of the Physical Society, seldom missing either meetings, and was at various times on the Councils of both; he was twice a Vice-President of the former.

He had abounding energy and great enthusiasm for all new developments in Optics, which was his chief love.

Rheinberg himself was a most likeable man, who will be greatly missed by a large circle of friends.

R. S. C.

MARK BENJAMIN

Mark Benjamin was killed in September 1941, when the aeroplane crashed in which he was returning from a visit to America. He was born in 1909, the eldest of a large working-class London family. He won a scholarship to the Regent Street Polytechnic, secured a high place in the L.C.C. open examination, and started work at County Hall when he was 18. But his heart was set on scientific work; finding no scope for it in his position, he abandoned it, and in 1928 took a very subordinate post at the Research Laboratories of the General Electric Co., Ltd. Then he studied at evening classes and took a London degree in Honours Physics. His abilities were soon recognized. He was attached to the Valve group, and worked first with A. L. Reimann and then with C. J. Smithells. When both of these had left, in 1938, he was put in charge of one section and became responsible for the co-operation of the whole group with the factory.

To great intellectual power, Benjamin added tireless energy and an irresistibly infectious enthusiasm. Tact he would not have claimed; but his criticism was always so constructive, and so evidently inspired only by a desire to achieve the common end, that it was tolerated, and even welcomed, almost everywhere. He was the ideal industrial scientist, combining the abstract intuition of the investigator with the practical sense of the manufacturer.

In his short career Benjamin collaborated in several papers of general interest. With H. P. Rooksby, he studied the thermionic emission from mixed barium and strontium oxides; they reached the conclusion that the function of the strontia is to form a solid solution with the baryta to which the monatomic layer of barium adheres more firmly than it does to pure baryta. This work won him the degree of Ph.D. With R. O. Jenkins he investigated the diffusion of barium on metal surfaces at room temperature; but here their conclusions were shown by Becker to be too general. His latest work, also with Jenkins, concerned the autoelectronic discharge from fine metal points, which can act as a powerful electronic microscope, revealing the minute structure of the surface from which it takes place. Their beautiful experiments, some still in the process of publication, greatly extended the work of Müller; in particular they produced evidence of the surface migrations of metal atoms under high electric fields at temperatures far below the melting point (e.g. tungsten at 1200° K.).

Benjamin leaves a widow, to whom his work owes much, and two young children.

N. R. C.

TOM HARRIS

T. Harris, who died on 15 January 1942, at the comparatively early age of 56, was born in London on 4 November 1885. He was educated at Upper Latymer School and later at the Royal College of Science, where he took a degree in Physics. From London he went to Cambridge as an advanced student and, although strictly a London man, it was always Cambridge that was his intellectual home. I could never be quite sure whether his attraction was in the scholarly atmosphere of the picturesque colleges, or in the inspiration that he derived from that great physicist, to whom so many owe so much, Sir J. J. Thomson.

On leaving Cambridge in 1911 Harris was appointed lecturer on the Physics Staff of Queen Mary College, London, but it was not very long before the war of 1914 claimed him for a commission in the Royal Engineers. He saw service in England and France—mostly in connection with meteorological work.

At the end of the war, Harris returned to Queen Mary College. He was an excellent colleague, and the years that immediately succeeded were, to me, and, I think, to him, of the pleasantest. It was in the discussion of physical problems that arose as a matter of general interest or out of research that Harris showed up so well. I always admired the way in which he would never rest content until he was quite clear about the problem under discussion. His mind was not of the type that would hurry a piece of research so as to be first in the field. He cared little for notice. To be clear himself and to explain clearly to others seemed to be his only object.

Harris, like myself, left Queen Mary College in January 1921, and joined the Research Dept. at Woolwich. Although we were now in widely separated branches, I used to see him frequently. It was here, of course, that the major part of his research work was done. It is a pity that its confidential nature makes discussion of it impossible, for Harris had a power of attacking a problem both analytically and experimentally, and I can remember the enthusiasm with which his assistants spoke of his powers.

Harris's work in the present war took him once again to Cambridge. Knowing his pleasant memories, I could not imagine more congenial surroundings, but, alas! I knew little about the trying nature of his work and of his failing powers.

His loss will be deeply felt by all who knew him.

He leaves a widow and a grown-up son and daughter.

G. D. W.

REVIEWS OF BOOKS

University Physics. Part III, *Light*, by F. C. CHAMPION. Pp. 172. (London and Glasgow: Blackie and Son, Ltd., 1941.) 5*s*. 6*d*. net.

The value of a text-book on Light can be very largely made or marred by the diagrams it contains. Optics is a notoriously dry subject for most students, but well-drawn and accurate diagrams can go a long way to lighten the algebra and make it intelligible. Unfortunately, the illustrations in this book have been very carelessly drawn. There are at least 24 diagrams that are incorrect in some way or another—rays refracted in impossible directions, incident and reflected rays on the same side of the normal, eyes placed at the wrong position to see the image, astigmatic images not shown at the correct intersections, object and image-points that could not possibly be conjugate to one another, innumerable interference fringes where only a few actually can be formed, and so on. All this makes the task of the student very difficult. It is not assisted by the hesitancy with which the author uses his sign convention, leading on p. 41 to the curvature term in the equation for the focal length being derived as $(1/r_1+1/r_2)$, while on p. 28 it had been derived as $(1/r_1-1/r_2)$—a discrepancy not to be attributed to the printer. The book as a whole is altogether too sketchy to be of real value, the subject matter which the author attempts to cover being far too wide for the 150 pages devoted to it. If there was a need for a new book on Light, it certainly has not been met in this volume.

W. D. W.

The Twin Marchant Calculating Machine and its application to survey problems. By L. J. COMRIE, Ph.D., M.A. Pp. 40. (London: Scientific Computing Service, Ltd., 1942.) 10*s*.

This large booklet is actually the third, but a largely rewritten and re-titled, edition of a work which appeared in 1938. The description applies not only to a twin but also to the single *Marchant* machine. The twin machine has been cleverly constructed, apparently with complete success, by connecting two singles back to front; the two main operating handles are replaced by a single crank acting through gearing on the two multiplying mechanisms. At the moment this is apparently the only twin machine available, though it is expected that the Swedish *Facit* machine will be obtainable after the war. It is noteworthy that the greater number of modern machines in use for general calculation are developments of the original design by Odhner, a Swede who lived in St. Petersburg; the *Leibniz bar*, in spite of its silent working, has evidently not been found capable of similar development.

With this and many other machines and a staff of highly trained assistants, Dr. Comrie is now in a position to undertake the computation of the most complex functions and the resolution of onerous problems, such as those of least squares with more than, say, 20 unknowns. These problems could not be described in a booklet of reasonable dimensions; accordingly no more than simple operations are considered. Particularly valuable are the hints given as to the proper use of the machines and the short cuts available. The twin machine is, of course, used especially for the common case of two or more multiplicands with a common multiplier.

J. T. McG.

Meteorological Instruments, by W. E. KNOWLES MIDDLETON. Pp. viii+213. (Toronto: The University of Toronto Press; London: Sir Humphrey Milford, at the Oxford University Press, 1941.) 11*s*. 6*d*.

No comprehensive English text on meteorological instruments has been published in the present century; we have had only articles in Glazebrook's *Dictionary* and the elementary though well-conceived *Meteorological Observer's Handbook*, now in a thirteenth edition. Kleinschmidt's *Handbuch der Meteorologische Instrumente* (1935) is the only thorough treatment available, and this will hardly lose its place with the appearance of the book under review, though the latter, much smaller in bulk, supplements Kleinschmidt in several practically important respects and has the merit of being more critical and less diffuse. Middleton's work is directed essentially to the meteorologist responsible for day-by-day observation of the common meteorological quantities, those which go to making the surface and upper-air synoptic weather charts. Thus he deals with the measurement of barometric pressure, air temperature, humidity, precipitation and evaporation, wind near the earth's surface and in the free atmosphere, cloud characteristics, sunshine duration and radio exploration of the free atmosphere. Less general but fundamental measurements on solar and terrestrial radiation, atmospheric opacity, etc., are not considered, but a short section is devoted to the measurement of cloud-particle size and water content.

There are excellent features to the book. Thus the author states clearly the principles of good meteorological instrument design, principles which many investigators learn only by long experience in field measurement. He makes a point of the accuracy of observation which is worth attaining, and that which any instrument does in fact give—as opposed to figures based on unconsciously idealized conditions. The treatment of a property of a class or sub-class of instruments is given point to by a few numerical figures; these help to clinch an argument, on instrument lag for example, and add much to the usefulness of the book. And one admires the balanced treatment of radio-sondes, which have been developed so rapidly in recent years that the balanced view is difficult to obtain. Also, as a Canadian meteorologist in close touch with British and American practice, the author claims to present the better features of both.

One's criticism of the book is that, within its professed limits, the treatment is not sufficiently thorough and up to date, and an addition to the two hundred pages should have been made to fill the lacunae. Hygrometry is an example. Meteorologists, indeed physicists generally, should be made aware of Simon's and of Powell's work showing how the reduction of wet-bulb size may almost eliminate the effect of ventilation on the psychrometric "constant". The hair hygrometer is well treated, but only passing reference is made to the superior qualities of gold-beater's skin, and for neither instrument is it explicitly stated that at sub-freezing temperatures the indications are of relative humidity with respect to super-cooled water and not to ice.

Other omissions relate to aneroid barometers—the desirability of using beryllium bronze to avoid elastic errors in aircraft and radio-sonde instruments; to radiation effects on meteorological instruments—no general treatment is attempted; and to the effect of Reynolds number on the drag coefficient of balloons.

The book shows evidence that the subject, from a period of relative stagnation, is once again very live. A meteorologist may now have before him in his working room indications or records of what is happening "out there" in the atmosphere, at least for the more important variables. Other notable advances are in balloon theodolite design, cup and pressure-plate anemometry, a recording dew-point hygrometer and the radio-sonde already noticed.

The text is concisely written, well indexed and referenced and attractively presented. There are few serious errors or misprints, but two may be noted. The error in temperature indication due to airspeed is given (p. 54) proportional to air density, while it is, in fact, independent of the latter ; and there is confusion in nomenclature in the treatment of balloon lift (p. 149). The experimental meteorologist will be glad to possess the book, though wishing for more, and the general physicist will find it of use when measurements of the qualities of an atmosphere are to be made. P. A. S.

Reports on Progress in Physics, Volume VIII (1941). General Editor, W. B. Mann. Pp. iv+372. (London : The Physical Society, 1942.) 25*s*.

This new volume contains articles on : *The teaching of theoretical physics in universities* (10 pages), by M. Born ; *The theory of the specific heat of solids* (20 pages), by M. Blackman ; *High-speed centrifuging* (18 pages), by J. W. Beams ; *Applications of spectroscopy to combustion* (21 pages), by A. G. Gaydon ; *Friction between solid bodies* (19 pages), by R. Schnurmann ; *The general physical constants* (45 pages), by R. T. Birge ; *Recent television developments* (51 pages), by V. K. Zworykin and R. E. Shelby ; *Instrumental technique in astrophysics* (14 pages), by A. Hunter ; *Dispersion in the far infra-red* (12 pages), by L. Kellner ; *The M.I.T. wavelength project* (19 pages), by G. R. Harrison ; *Molecular electronic spectra, dispersion and polarization* (43 pages), by R. S. Mulliken and C. A. Rieke ; *Symmetry properties of nuclear levels* (44 pages), by E. P. Wigner and E. Feenberg ; *Photometry* (20 pages), by H. Buckley ; *Discharge phenomena in gases* (30 pages), by R. W. Lunt, A. von Engel and J. M. Meek. An author index (5 pages) is provided.

With Volume VIII is issued the List of References for the *Report* by F. Eirich on " Viscosity of Substances of High Molecular Weight in Solution " which appeared in Volume VII (1940).

RECENT REPORTS AND CATALOGUES

Rotameters. Pp. 5. Rotameter Manufacturing Co., Ltd., Vale Road, Portslade, Sussex.

Kryptoscreen X-Ray Paper and Brytex Intensifying Leaf. Pp. 6. Ilford, Ltd., Ilford, London.

The Cambridge Bulletin (No. 88, Spring 1942). Pp. iii+16. Cambridge University Press, Bentley House, 200 Euston Road, London, N.W. 1.

THE PROCEEDINGS OF THE PHYSICAL SOCIETY

VOL. 54, PART 4 1 July 1942 No. 304

RESEARCH ON COLOUR PHYSICS AT SOUTH KENSINGTON, 1877–1942

BY W. D. WRIGHT, A.R.C.S., D.Sc.,
Technical Optics Section,
Imperial College of Science and Technology

Chairman's Address to the Colour Group, delivered 11 *February* 1942

ABSTRACT. The colour research carried out by Sir William Abney at South Kensington is briefly reviewed and compared with that carried out more recently in the Technical Optics Department of the Imperial College. The instrumental methods employed by Abney and the writer are examined in relation to current problems in the design of colorimeters and spectrophotometers, including mechanical construction, viewing of the matching field, intensity control, and precision of the optical parts. The spectral-mixture curves measured at the College are discussed in connection with the variation of colour matching amongst different observers and with the possibility of correcting these variations. The work on colour discrimination is reviewed, and its bearing on the specification of colour tolerances and the derivation of a uniform-chromaticity scale discussed. Suggestions are put forward for future research on colour physics.

§ 1. INTRODUCTION

It must be rare that almost continuous research on one subject at one place has been maintained for so long a period as 65 years, yet it was in 1877 that Sir William Abney first entered the Science and Art Department at South Kensington. After only a few months' residence there, Abney established a laboratory in a hideous iron building which, with its companions, became known as the "Brompton Boilers". Ultimately, when various colleges were incorporated in one organization, namely, the Imperial College of Science and Technology, Abney acquired excellent laboratory facilities in the new building of the Royal College of Science, but this was not until thirty years later, in 1907. A great deal of Abney's work was on photography, and it was through photography that his interest in the spectrum and in colour vision was aroused. He continued active experimental work up to the outbreak of the last war, having carried out pioneer researches on colour vision and colour measurement.

It is a remarkable fact that not only has colour research been continuous for 65 years, but some of the problems that have been tackled in the last few years are almost identical with those investigated by Abney. The main hope that we have of finding solutions where Abney failed lies in the fact that in the intervening years many misconceptions about colour have been removed, a great deal more has been learned about the physiological processes of vision, and the technical equipment available to workers in photometry and allied subjects has been vastly improved.

In his later days, Abney was assisted in much of his work by Professor W. Watson, but little was done on colour during the last war until the Technical Optics Department was formed in 1917. Abney's colour-patch apparatus was then included as part of the equipment of that department, and research and instruction were continued by Professor L. C. Martin and others. In 1926 the Medical Research Council, through the kindly influence of Sir John Parsons, became the fairy godmother of colour research at the College, and they have held that office ever since. I should like to take this opportunity of publicly acknowledging our very real indebtedness to them for the most generous help we have received and for the complete freedom we have been given to develop research along the lines that appeared to us to be most profitable.

A programme of work was prepared by Professor Martin and Professor H. Hartridge, the first problem tackled in 1926 being the re-determination of the spectral-mixture curves, information that was, as reported by Mr. Guild at the Optical Convention in 1926, urgently needed for colorimetric purposes and was also of very considerable physiological interest. I was invited to undertake this investigation under the supervision of Professor Martin, and I have a further debt to acknowledge for the help and support which Professor Martin has always given me, and again for freedom in developing the research. My first task was to design a new trichromatic colorimeter, and I was fortunate enough to adopt a general arrangement that was found subsequently to lend itself to many different types of observation beyond the original problem of the colour-mixture curves. Research with that apparatus has been in progress continuously since 1926, although I have not myself been at South Kensington during the whole of that period.

So much for a brief historical outline of the activities here from 1877 to 1942. I propose now to deal with some of the salient points of the actual investigations and their impact on current problems in colorimetry, and to contrast the methods adopted by Abney with those used more recently in the Technical Optics Department.

§ 2. INSTRUMENTAL ARRANGEMENTS

Abney's colour-patch apparatus is too well known to require detailed description here, and those who wish for more information can readily obtain it from Abney's original reports (Abney, 1913). Similarly, I do not wish to spend time on the details of my own colorimeter (Wright, 1927, 1939), but rather to compare certain features of the two instruments and discuss them in relation to modern instrument design.

Consider, first, the method of intensity control that has been used. In each case it was necessary, in order to produce a colour match, to vary the intensities of three beams taken from different parts of the spectrum. Intensity control is a perennial source of trouble to designers of photometric, and especially spectrophotometric and colorimetric, apparatus; in a trichromatic colorimeter the problems may be particularly awkward because whatever device is chosen has to be used in triplicate, and space may prove a serious restriction on the choice that can be made. Consider the possibilities: (1) the inverse-square law, (2) polarization devices, (3) rotating sectors, (4) variable apertures, (5) photometer wedges, (6) grids.

(1) is excellent when the illumination of a screen is to be varied by varying the distance of a lamp from the screen, but it is usually quite impossible to employ it in the type of apparatus we are considering. (3) is certainly possible in some instruments, and may consist of a series of fixed-aperture sectors or a variable sector, e.g. the type of sector used by Abney in some of his experiments. Fixed-aperture sectors provide probably the most reliable method we have of cutting down the intensity of a beam in some given ratio, and as one stage in the control of the intensity in a spectrophotometer, e.g. in Guild's spectrophotometer (Guild, 1924), it is excellent. But its use in triplicate in a colorimeter is out of the question. Use of (6) may be a possibility in some instruments, when the grids are placed in beams of wide and uniform cross-section, but these are not the conditions that usually obtain in spectrometric apparatus, and in any case the accurate production of fine grids is no easy matter and, so far as I am aware, they are not used extensively.

There are thus three methods remaining, namely, polarization devices, variable apertures, and photometer wedges, and all three have, in fact, been employed in trichromatic colorimeters. Abney controlled his intensities by varying the width of slits placed in the plane of the spectrum; this is, in effect, a variable-aperture method of intensity control. So, too, is the device used in the Donaldson colorimeter (Donaldson, 1935), although, because the areas involved are so much larger, the control is more accurate than that used by Abney. In the Guild colorimeter (Guild, 1925), the variable-area method is also used, but the method is related to the rotating-sector device through the action of the rotating prism. In the Helmholtz colorimeter, which was employed so extensively by König (1903), polarizing prisms are incorporated to vary the intensity, while in my own instrument I have used photometer wedges.

Personally, I have a prejudice against polarization methods in spectrophotometric and colorimetric apparatus, a prejudice which is not, however, shared by everyone; they are undoubtedly easy to operate and possess little irregularity in the intensity-rotation relation, but they are wasteful of light, they are subject to unsuspected sources of error from partial polarization that may have taken place in other parts of the optical system and, without complicating the system by the addition of further polarizing prisms, the intensity scale is very condensed at the low-intensity end. Photometer wedges are very convenient from almost every point of view except that they may not be strictly neutral and, although the intensity variation is closely linear when plotted as density against length of wedge, very exact linearity is only possible with great care in their manufacture. Variation of aperture is also very convenient but, as evidenced by the care required in the Donaldson colorimeter to make special correcting templates, the method cannot automatically be assumed to give accurate intensity control without thorough testing. Certainly, in Abney's case, the area was so small that the accuracy must have been strictly limited, and yet it must have proved a very handy method of control. There does still seem to be an opportunity for someone of an inventive turn of mind to design a simply operated intensity control, with an open-scale, very smooth calibration curve so that it can be read to a very high accuracy and with neutral transmission with respect to wavelength.

One of the big disadvantages of Abney's apparatus was the rather low intensity of his matching field. This was inevitable because his method of mixing his three primaries was to focus the three pencils on to a diffusing screen. The reduction in brightness compared to that seen when the eye is placed immediately behind the slit in the spectrum is very considerable, and yet the use of a diffusing surface has much to commend it. I certainly envy the ease with which Abney could demonstrate the way in which his three primary beams were being mixed; he had only to hold a card behind his three spectrum slits and gradually withdraw it, when the three beams would be seen first individually and then gradually overlapping. There would be no doubt at all what was going on. Nowadays, nearly all observations are made by looking through a tube—we hardly think it scientific unless we do so—and unless we happen to be familiar with the apparatus, there is usually a good deal of mystery about what is going on inside. I am not advocating that we should forthwith scrap the Maxwellian field of view—that is generally both undesirable and impossible—but there may be occasions when we might dispense with it with advantage, more especially, I believe, in apparatus which we are using for instructional purposes, and particularly instruction in general principles.

When the matching field is projected on a screen, the actual act of observation is very much easier, and this is all to the good in so far as it adds to the comfort of the observer. There is also the advantage that both eyes can be used, as they are, of course, used in everyday observations. But for some observations there is the disadvantage that the diameter of the eye-pupil is uncertain, and hence the amount of light falling on the retina is indefinite. Viewing the field through the exit-pupil of an instrument to some extent overcomes this trouble, but it does not do so completely unless some method is adopted whereby the eye-pupil and instrument-pupil are centred relative to one another and maintained in that position. For, some years ago, Stiles and Crawford (1933) showed that a narrow pencil of light passing centrally through the pupil of the eye causes a greater luminous effect than the same pencil passing near the edge of the pupil. If the exit-pupil of the instrument we are using is fairly small, the apparent brightness of the field will vary as the observer's head moves. This can be prevented most effectively by fixing the head with the dental-impression method used by Stiles and Crawford, and I have, in my colorimeter, provided a mounting to which each observer can clamp his own dental impression and thereby centre and fix his head. Although I first incorporated this for some binocular matching experiments, we now use it for all observations with the colorimeter as a matter of routine, and it is my belief that the consistency of our results has been improved through its use. Certainly, it prevents the appearance of coloured fringes round the field of view, fringes that are sometimes rather disconcerting to the observer, especially the inexperienced colour matcher, and I would urge instrument manufacturers to consider whether it might not prove advisable to use some such fixation device in certain types of visual instrument, particularly those in which the exit-pupil is small or unsymmetrical, e.g. a slit.

Another feature of instrument design that I would like to take up with manufacturers is the general question of size. Because of the need to keep the costs of instruments down to a minimum, manufacturers tend to make their instruments as small as is consistent with satisfactory operation, and if they have made a correct estimate of the economic position, I cannot, of course, quarrel with their decision. But I suspect that in not a few cases they make the instruments small because they like them that way. They are proud of the intricacies of a complicated mechanism and of the compactness of a design. For the research worker, I do not believe this is a correct view to take, and I would like to make a plea for larger instruments, where cost and space permit. As a small example of two equivalent optical systems, consider the optical arrangement at the eyepiece end of my colorimeter. The method I have used is shown in figure 1 (*a*); an alternative arrangement is shown in figure 1 (*b*). Each would bring the light to a focus in the plane of the exit-pupil and give the same-sized patch of light in that plane, the photometer prism would be seen in sharp focus and under the same visual angle for an eye focused at infinity, yet the second arrangement is shorter than the first by nearly the focal length of the lens L_1, about 10 inches in my instrument. Yet I have found that extra length very useful, as it has provided me with space for other devices that would have been very difficult to fit in otherwise. In research work, it is impossible to know in the initial stages all the types of observation that may ultimately be made with any given piece of apparatus, and if the manufacturer leaves no space for additions and modifications, then the instrument is to that extent less satisfactory as a research tool.

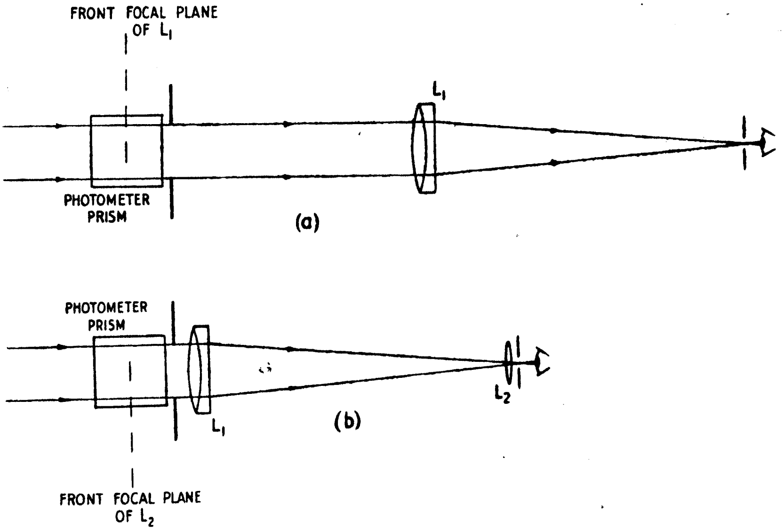


Figure 1.

However that may be, both Abney's instrument and my own have been on a large scale. My original model was a somewhat clumsy apparatus of glass and brass and wood, spreadeagled over a considerable area of benches and tables. In a later model I have built the base from a central cast-iron pillar and heavy cast-iron arms radiating from the pillar. Although the weight of the optical parts to be supported is quite small, it is important that the various optical parts, situated, as they are, several feet apart, should not move relative to one another. The arrangement I adopted has proved very satisfactory both from the point of view of rigidity and accessibility and was a great improvement over the earlier model spread over a large flat area. In an instrument costing an amount of the order of £500, I consider £50 a reasonable fraction to spend on a heavy base of this type; that has been my experience, and I mention it because it may be of some interest to others in various research laboratories. The question of cost will, however, take on a different appearance where a complete instrument has to be ordered from a firm, instead of getting the various components made in your own workshop or by small firms outside, and assembling the parts in the laboratory.

I have suffered a good deal of ridicule from my more mechanically-minded friends over my extensive use of strings, weights and pulleys as a means of moving my photometer wedges for colour matching and moving the test-colour prism for hue-discrimination measurements. It would no doubt be possible to design a rack and pinion for each wedge, operated through a flexible cable or by rods linked with universal joints, but I do not believe such a method would give so smooth or easy a movement as that given by a drum with a string wound round in one direction leading to the wedge and in the other direction to a counter-balance weight. Further, in the case of the test-colour prism, this may sometimes have to be moved only a millimetre by one observer in making a just noticeable step, while a colour-blind observer may need 2 or 3 centimetres for a step; or again, the person operating the apparatus may wish to transfer the prism from one end of the spectrum to the other in the space of a few seconds. I do not think these requirements could be met by any rack-and-pinion device without a great complication of mechanical design. I have been told that instrument firms would hesitate before putting an instrument on the market that was operated by strings and pulleys; I would venture to doubt their judgment on this point, if it is a true statement of their views. I would certainly claim that no colorimeter on the market has the three knobs that control the amounts of the primaries in such a convenient position as I have been able to arrange by the aid of my string system.

There is one last point I would like to make on the question of instrument design. I would like to emphasize the advantages that can sometimes be obtained by making use of the high precision that is possible in the manufacture of prisms and lenses. When I built my first instrument, I had not had sufficient experience to realize what might be gained in this way, and as a result, I provided a number of adjustments of the rotating and tilting kind that could have been avoided by improved optical design. In the later model, I incorporated a number of alterations which would take too long to describe in detail, but which included tall dispersing prisms instead of two shorter ones mounted above one another, a front-surface aluminized mirror with a rectangular hole in it to provide two strips of mirror in one plane, instead of two independently mounted reflecting prisms, a photometer prism with its angles made to a high degree of accuracy, so that two beams which are parallel before they enter the prism are parallel when they leave it, and so on. By means such as these, the new colorimeter was almost automatically in adjustment once the components were assembled in position, and an adjustment that had previously been very awkward, namely, the bringing of the images of the two spectra to a position of coincidence in the plane of the exit-pupil, now followed inevitably from ordinary optical principles, within the limits set by the flatness of the prism surfaces and the mirror surface, and by the aberrations of the corrected lenses employed. Optical instrument makers as a whole naturally appreciate the value of optical tricks of this kind, but I believe there is a tendency among those concerned with the assembly of optical apparatus for experimental and research purposes to provide elaborate and unnecessary mechanical adjustments that make the handling of such apparatus far more difficult and irritating than it need be.

§ 3. THE COLOUR-MIXTURE CURVES

Both Abney and myself were concerned with the measurement of the amounts of the red, green and blue instrument primaries that were needed to match the various colours through the spectrum. Data of this kind are fundamental to any method of colour specification, and the data we recorded at the Imperial College were combined with those obtained by Guild at the National Physical Laboratory to form the basis of the colour-matching characteristics of the normal observer subsequently adopted by international agreement. That is now past history, and there is no need to go over the ground again here. One aspect of these measurements is, however, still of current interest and practical importance, namely, the variation of the colour-matching properties among a group of observers, their differences from the standard observer, and the possibility of correcting small departures from normality so that all measurements would be standard measurements. I would therefore like to discuss our experiments from that point of view, and must first refer to the system of units that I adopted in recording the colour-mixture curves.

When specifying a colour on the trichromatic system, it is not only necessary to define the primaries themselves, e.g. by their wavelength if they are monochromatic colours, but also to define their relative magnitude. It is common practice to do this by making a match on some specified white and adjusting the units so that equal quantities of the primaries are required in the white match. This is perfectly sound practice for colorimetry, when the main aim is to obtain a result as near as possible to that which would be obtained by a normal observer and for which a measuring technique is required that tends to remove the effects of individual differences of colour vision. My interest, on the other hand, was in part to investigate those differences as well as to obtain a mean set of curves, and the use of white to define the primary units has one disadvantage for this purpose. Over a retinal area of some 3° in the foveal and macular regions, there exists a yellow macular pigment that varies in density from one observer to another, and before any light can reach the fovea, it has to pass through this pigment. By the time, therefore, that a pencil of white light of some standard energy distribution reaches the light-sensitive part of the retina, it has become more or less yellow and can no longer be regarded as a standard white. On the other hand, a monochromatic radiation from only a narrow band of the spectrum could not change in energy distribution as it passed through the yellow pigment, but would merely suffer a reduction in intensity. Yet if the primary units were based on a white match, the white would always be recorded as having the same colour coefficients whatever the density of the pigment, whilst the corrected coefficients of the spectral colours would vary with its density. Such variations would be erroneously put down to differences in the light-sensitive mechanism itself, a misleading conclusion if we hoped to use the data to correct for individual differences in colour vision. The alternative that I used was to base the units on matches on two monochromatic radiations, a yellow at $0{\cdot}5825\ \mu$ for which the red and green units were made equal, and a blue-green at $0{\cdot}4940\ \mu$ for which the green and blue units were equalized. These unit colours were unchanged in colour after their passage through the yellow pigment, hence any differences in the coefficients for the spectral colours that were recorded for different observers could not be due to pigment differences, but were, presumably, due to variations in the response mechanism itself. On the other hand, the effect of pigment variations was all thrown on to the coefficients of the white point, and the position of the white point in the colour chart was, in part at least, an indication of the colour of the macular pigment.

The results of the white-point measurements (Wright, 1928) suggested very strongly that much of the observer variation is due to pigment differences, i.e. considerable variations in its density and slight differences in dominant hue. I have never made any attempt to correct for these differences by means of a yellow or blue filter, but this should certainly be possible, just as correcting filters have been suggested for luminosity measurements (Ives and Kingsbury, 1915). The density of the correcting filter would, of course, have to be determined for each individual by a separate white match. It is generally considered that the density of the pigment increases with the age of the observer, and this is confirmed by my own experience, as my white point has moved during the past 15 years from a position well on the blue side of the average to a point slightly on the yellow side. But it is also true that the pigmentation varies considerably from one individual to another of the same age, so that while each individual may get yellower with age, the starting points will be different for different people. For example, the yellowest point that I obtained from a group of 35 observers was for an observer who was only 20 years of age.

It is difficult to feel so happy about the possibility of correcting for the variations in the colour-response mechanism that are revealed by the differences in the spectral-coefficient curves. It has already been implied that the employment of a colour filter would have no effect on the coefficients of spectral colours, and would not, therefore, achieve the correction required. The most promising line of attack would probably be to find the magnitude of the errors in various parts of the colour diagram and apply suitable corrections; even this would not be completely satisfactory, as the amount of the correction would not be unique for any given point in the colour diagram, since any one colour can be produced by many different energy distributions. Complete elimination of the observer trouble is therefore hardly to be hoped for, and the adoption of substitution methods to reduce the errors to second-order magnitudes should be followed in visual colorimetry.

It is interesting to compare the spectral-coefficient curves obtained by Abney with those just described, when both sets of data are referred to the same primaries and the same units. It then appears (Wright, 1928) that there is surprisingly little difference between the two, so that the doubts about the numerical accuracy of Abney's results are not confirmed. This is very gratifying; the major difference is in the position of his white point, and this might well be expected, since he made no pretence of using a white with a precisely determined colour temperature.

§ 4. COLOUR DISCRIMINATION

One type of observation to which we have devoted a good deal of time is the measurement of the just-noticeable colour difference with change of wavelength or saturation or position across the colour chart. Professor Martin, with the assistance of Mr. Warburton, was the first to carry out such investigations at the Imperial College, in this case on saturation discrimination (1933); since then, data have been recorded on hue discrimination for both normal and colour-blind observers and, quite recently, for the whole area of the C.I.E. colour chart. One of the weaknesses of Abney's work was that he made no observations on the sensitivity of the eye to colour differences; his colour-patch apparatus, unfortunately, did not lend itself to that kind of observation. As a result his conclusions about colour deficiencies were rather incomplete, since weakness in colour discrimination is the most outstanding, and in practice the most serious, feature of colour blindness.

I have only time to deal with the last of our investigations on discrimination (Wright, 1941), which had, as its immediate object, the testing of the standard C.I.E. colour chart (Smith and Guild, 1931) to find the extent to which equal colour differences are represented by equal lengths in the chart. Previous work elsewhere had indicated that the departure from equality would be considerable; but I believe this was the first experimental check of practically the whole area of the chart. We found that equivalent steps in the blue and purple regions of the chart could be represented by quite short lines, but in the green corner the lines were some 20 times longer. This makes the C.I.E. chart an inconvenient diagram for the representation of colour tolerances and for judging the significance of colour differences; the ideal diagram would be one in which a colour tolerance could be specified merely by the radius, r, of a circle drawn about a point C in the diagram, C representing the specified colour and the circle representing the area within which any attempted match of C must lie if it is to be acceptable. There has been much activity in the United States in the last few years aiming at the discovery of an oblique projection of the C.I.E. chart that would give a more uniform distribution of the colours, and when our results were expressed on the Uniform Chromaticity Scale of Breckenridge and Schaub (1939), it was apparent that the steps were very much more nearly equal in size than before, although they still varied in some parts of the chart by as much as 4 : 1. Other data are being recorded in the U.S.A., and when they are all collated, it will be interesting to see whether an improved projection can be found in which the steps are more nearly equal over the complete area. Exact equality is certainly not to be expected. If the suggestion for a change from the present C.I.E.

system as a basis for standard colour specifications is made, the existence of the Colour Group should ensure that the pros and cons will be fully discussed in this country, and in the meantime, it may be advantageous to provide further opportunities for the discussion of the trichromatic colour diagram as a medium for the specification of colour tolerances and the extent to which it is being used for this purpose in practice.

MacAdam (1939) has maintained that tolerances are so often governed by manufacturing considerations, and in any case vary according to the type and conditions of observation, that no single chart could serve as a uniform scale for all cases. For this reason he considers that the C.I.E. chart should be retained. In a recent paper (Wright, 1941) I ventured to doubt whether the conditions of observation would cause any significant change in the relative size of the steps, and whether this could be used as an argument against a change of system. Since then, however, I have had occasion to make some hue-discrimination observations when the two comparison fields consisted of very small circular apertures separated by an angle of 1°. The angular size of the geometrical images of the two apertures was only 6′, but this was enlarged somewhat by diffraction, especially as they were viewed through an exit-pupil of 1 mm. The results showed that, in comparison with those obtained with a 2° field, the size of step as measured by wavelength difference increased in the blue-green to a greater extent than in the orange, the relative change being by a factor of the order of 2 : 1. As a matter of interest, this supports current practice in the choice of red as the ideal signal colour and the absence of both blue and green together in any one signalling system. The change in observing conditions is, however, rather extreme and might not be considered to affect the main issue about the merits of a change in the C.I.E. system, since small point sources are not typical test objects and are of no concern, for instance, to the dyeing or paint industry. Yet the effect is of some interest, for at the present time the specification of the colour of signal glasses and their tolerances is the most outstanding example of colour limits being expressed on the trichromatic system and the one most frequently used for illustration purposes in papers on the subject. To this extent, therefore, MacAdam can find support for his contentions, and there may, of course, be other examples of the same nature. Obviously the subject will be a fruitful source of discussion in the future.

§ 5. SUBJECTS FOR FUTURE RESEARCH

That gives me the cue for the final remarks I wish to make. What are the most fruitful lines of research on colour to follow during, say, the next 10 years and, since this is an account of work at the Imperial College, what problems should we seek to solve here? I am not now thinking of problems in colour vision; I have made no reference to the researches we have carried out that have had primarily a physiological bias, and there is little difficulty in planning a programme of work on visual subjects that would last for several decades. But in the physics of colour there may be some suggestion that most of the work has been accomplished with the adoption and general acceptance of the characteristics of the standard observer, although this is not my own view. There are certain matters to be tidied up, such as the tolerance question to which I have already referred, and other matters, such as the colour-matching characteristics of the extra-fovea, which should be of some colorimetric interest in connection with the use of large matching fields. These, however, are comparatively minor problems compared to the subject which seems certain to occupy our attention for many years to come, namely, spectrophotometry, and, in particular, photo-electric spectrophotometry. The United States are a long way ahead of us in this field, and the extensive use they are making of the method is an indication of its importance in industry. It is not as a means of colour measurement that its chief advantage lies, but rather as a method of following the effects of subtractive colour mixture. This includes most of the industrial applications of colour—dyeing, painting, printing, colour photography and so on; innumerable technical problems must arise in these industries, and in addition, there are highly academic problems of molecular structure, the relation between the chemical constitution of a substance and its absorption curve, the manner in which colouring materials are associated with the substances they colour, all subjects that call for close collaboration between physicist and chemist. I am hoping that we shall be able to help with the physics of these problems. There are ample problems to go round among the various organizations that have a legitimate interest in them, although I anticipate there will be a good deal of competition to acquire the first recording spectrophotometer that becomes available in this country, a competition for which I hope the Imperial College will at least be able to enter.

In preparing this lecture, I have been fortunate in being able to draw on material that has accumulated for over 65 years and inevitably I have had to restrict my remarks to a few isolated aspects of the work. I imagine, however, that nearly everyone in the Colour Group could prepare an interesting account of his own experiences from which we could all learn something, and it is my sincere hope that, as occasion arises, members of the Group will come forward and assist us in that way.

REFERENCES

ABNEY, W. DE W., 1913. *Researches in Colour Vision* (London: Longmans, Green and Co.).

BRECKENRIDGE, F. C. and SCHAUB, W. R., 1939. *J. Opt. Soc. Amer.* **29**, 370.

DONALDSON, R., 1935. *Proc. Phys. Soc.* **47**, 1068.

GUILD, J., 1924–5. *Trans. Opt. Soc.* **26**, 74.

GUILD, J., 1925–6. *Trans. Opt. Soc.* **27**, 106.

IVES, H. E. and KINGSBURY, E. F., 1915. *Trans. Illum. Engng Soc.* **10**, 259.

KÖNIG, A., 1903. *Physiologischen Optik* (Leipzig).

MACADAM, D. L., 1939. *Amer. J. Psychol.* **52**, 412.

MARTIN, L. C., WARBURTON, F. L. and MORGAN, W. J., 1933. *Med. Res. Coun., Spec. Rep. Ser.*, no. 188.

SMITH, T. and GUILD, J., 1931–2. *Trans. Opt. Soc.* **33**, 73.

STILES, W. S. and CRAWFORD, B. H., 1933. *Proc. Roy. Soc.* B, **113**, 496.

WRIGHT, W. D., 1927–8. *Trans. Opt. Soc.* **29**, 225; 1928–9. *Ibid.* **30**, 141; 1939. *J. Sci. Instrum.* **16**, 10; 1941. *Proc. Phys. Soc.* **53**, 93.

X-RAY STUDY OF CRYSTAL DYNAMICS:
AN HISTORICAL AND CRITICAL SURVEY OF EXPERIMENT AND THEORY

By KATHLEEN LONSDALE, D.Sc.,
Davy-Faraday Laboratory, London

Lecture delivered 16 *and* 21 *January* 1942 *in London and Cambridge respectively*

ABSTRACT. An outline is given of the reciprocal-lattice method of relating x-ray reflections to the density distribution in the crystal. The experimental work of Laval, Preston, Zachariasen, Raman and Nilakantan, Lonsdale and Smith, and others, is described in some detail, and it is shown that a spreading of the reflecting power outside the reciprocal-lattice points is proved. The main features of this extension of reflecting power are (1) that it is markedly dependent upon temperature, (2) that the surfaces of equal reflecting power are not spherical, (3) that the intensity of extra reflection is weak even in the neighbourhood of the reciprocal-lattice points, and falls off rapidly with distance from the lattice points. The simple diffraction theory, the Faxén-Waller theory and the Raman theory are briefly sketched and related to recent experimental work, and the necessity for distinguishing effects due to dynamic and static disorder of the crystal structure is emphasized. A description is given of the anomalous effects found for diamond, and finally it is shown that only the Faxén-Waller theory, which explains the spread of reflecting power as due, in the main, to the thermal (elastic) vibrations of the atoms and molecules, can account satisfactorily for the results observed for crystals of all kinds, and especially for single crystals of metals. The limitations of the theory, especially in respect to its assumption of small amplitudes of vibration, are indicated, and possible future developments are outlined.

---

§ 1. INTRODUCTION

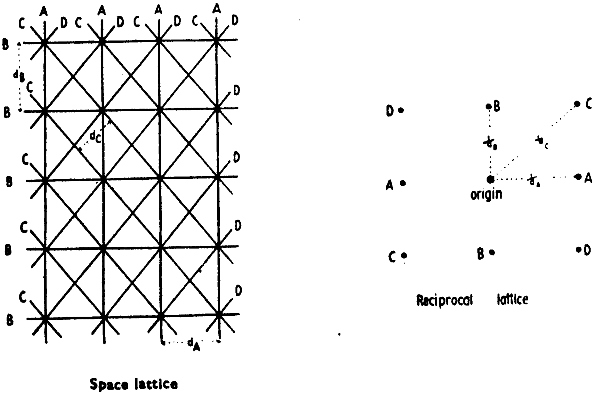
THE sharply defined spots on any x-ray crystal photograph will, if properly interpreted, give a partial or complete picture of the crystal as a three-dimensional diffraction grating composed of a regular arrangement of stationary atoms. It is well known, however, that the atoms are not stationary but oscillating, and that the picture, therefore, gives only the mean atomic positions. The dynamics of crystal lattices have been studied theoretically by Born and others (1912, 1913, 1933), and the x-ray interference effects due to the heat motions of the atoms were the subject of an early theoretical investigation by Debye (1914), who predicted a reduction of intensity of all normal reflections and a general background scattering independent of crystal orientation. These conclusions were supposedly confirmed by subsequent experimental observations (Jauncey, 1922 ; Jauncey and May, 1924 ; Woo, 1931; Coven, 1932; etc.) which we now know were made under unsuitable conditions.

When the intensities of x-ray reflection from the planes of the crystal lattice are measured and the structure factors calculated for purposes of structure analysis, allowance is always made for a reduction in intensity due to the heat movements, and it has long been realized that even after such corrections have been made, the electron-density map given by the most refined Fourier analysis of the crystal structure is blurred by these heat movements, since not only has each measured term of the Fourier series been reduced by a factor which can be expressed as an exponential function of temperature, but some of the later terms of the series have been thereby reduced below the measurable limit. For a really sharp and accurate Fourier contour map of the crystal, measurements of intensity would have to be made at very low temperatures, using x rays of various wave-lengths down to the shortest obtainable (Brill, Grimm, Hermann and Peters, 1939).

Very well-exposed photographs sometimes, however, show spots and streaks which are not part of the normal diffraction pattern at all. These have received various names to distinguish them from the Bragg spots (due to normal reflection by x rays of wave-length characteristic of the target used) and the Laue spots (due to normal reflection by "white" x radiation not characteristic of the target used). They have been called *diffuse*, *anomalous*, *extra*, *subsidiary*, *background*, *temperature*, *modified*, *dynamic*, *quantum*, *non-Laue* and *associated* reflections. (Research workers have occasionally referred to these reflections as *Preston* or *Raman* reflections. For reasons which will be made clear later in this survey, neither of these names has been generally adopted.) Since not all the observed effects are of similar origin, the most suitable term to use here will be *extra reflections*.

§ 2. THE RECIPROCAL LATTICE

In order to relate the observed x-ray diffraction effects to the static structure of the crystal, the use of the mathematical concept of "reciprocal space" (Ewald, 1921; Bernal, 1926) is helpful. In reciprocal space each set of infinite



(a) (b)

Figure 1.

regular planes, of spacing *d*, in the real crystal is represented by a point (see figure 1), whose distance from the origin is $1/d$ and whose radius vector lies along the direction of the normal to the planes. The *n*th order planes, of spacing d/n, will be represented by points along the same radius vector, but at *n* times the distance (that is, at n/d) from the origin. A development of this representation leads to the result that, corresponding to all the orders of all the sets of infinite, regular planes in the real crystal, there will be a regular, infinite lattice of points in reciprocal space.

The usefulness of the reciprocal-lattice idea is most apparent when an x-ray photograph has to be interpreted. The fact is, that any x-ray photograph resembles the reciprocal lattice of the crystal much more closely than it does the real crystal. For instance, each spot (point) on the photograph corresponds to a set of planes in the real crystal, but to a point in the reciprocal lattice. The nearer the spot is to the origin (that is, to the trace of the undeviated incident beam) the larger is the spacing of the corresponding set of planes, while the higher the order of reflection, the farther is the corresponding spot on the photograph from the origin. These reciprocal relationships are the result of the Bragg law of reflection $\lambda = 2 . d/n . \sin\theta$, according to which $\sin\theta \propto n/d$ for radiation of a given wave-length λ. The Bragg law is given a very simple geometrical interpretation in terms of reciprocal space (figures 2 and 3). Through the origin O of reciprocal space let a sphere be drawn, whose radius is $1/\lambda$ and whose centre C is at a distance $1/\lambda$ from O in an arbitrary direction. Suppose that some other reciprocal lattice point P lies also on the sphere. If $OP = n/d$, and the angle OCP be called 2θ, then it is obvious that

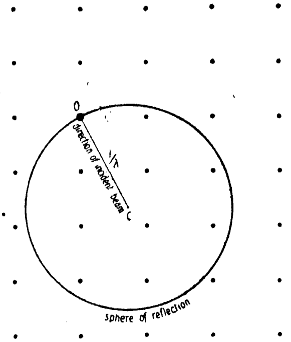


Figure 2.

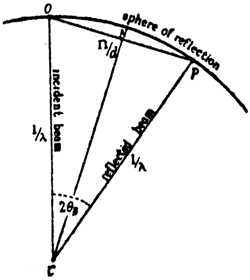


Figure 3.

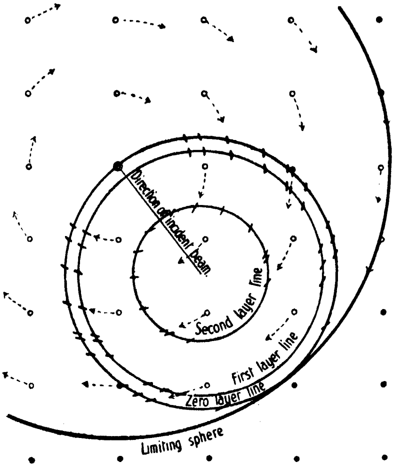
$$\sin \mathrm{OCN} = \sin \mathrm{PCN} = \sin\theta = n\lambda/2d,$$

and CN, the normal to OP, must be parallel to the set of planes reciprocal to P. Thus any point P which lies on the sphere may be said to obey Bragg's law, the direction CO being that of the incident beam and CP that of the reflected beam. The sphere itself is known as the sphere of reflection, and each point in the reciprocal lattice has an intrinsic reflecting power depending upon the structure of the crystal.

If λ and d are comparable, then for any arbitrary direction of the incident beam CO, the sphere of reflection will not *in general* pass through any points of the reciprocal lattice other than the origin (figure 2). In other words, a monochromatic x-ray photograph of a crystal which is arbitrarily set in any stationary position will in general be quite blank except for the trace of the undeviated incident beam. Various devices are in practice adopted in order to obtain reflections from a crystal; some of these are now described.

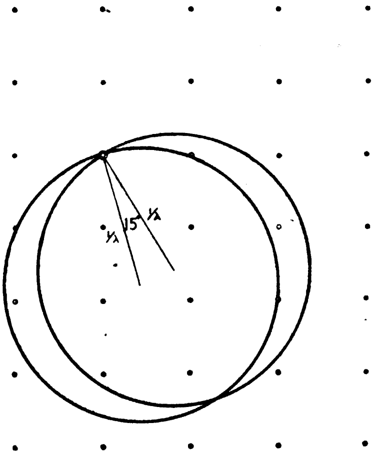
Methods for use with characteristic or monochromatized radiation; (rotation, oscillation or powder methods)

(1) The crystal may be *rotated* about an axis normal to the incident beam. Each point in the reciprocal lattice traces out a circular path about this axis, and in doing so intersects the sphere of reflection twice in each complete rotation (providing, of course, that it lies within the limiting sphere, for which $\sin\theta = 1$). If the direction of the incident beam lies in one of the main nets, or layers, of the reciprocal lattice, then all the reflecting points in that net intersect the sphere on the equator. Successive nets of the reciprocal lattice intersect the sphere in successive parallel small circles. These, projected on to



○ points which reflect during rotation.
• points outside limiting sphere, which cannot reflect.

Figure 4. Rotation method.

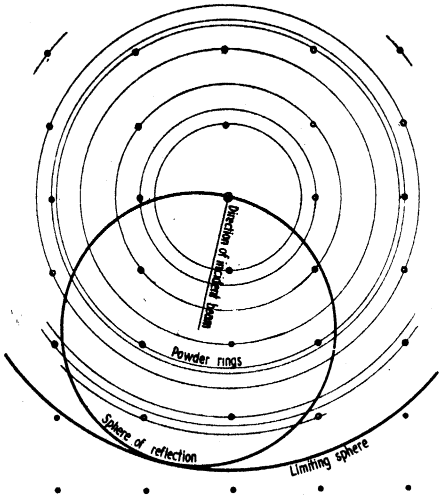


○ points reflecting during oscillation of 15°.
• points not reflecting during oscillation.

Figure 5. Oscillation method.

a cylinder or a plane placed parallel to the axis of rotation, give the layer-lines typical of such rotation photographs. In the case of a plane photograph it is not, of course, necessary for the plate or film to be perpendicular to the incident beam.

(2) The crystal may be *oscillated* through a known angle about an axis normal to the incident beam. The oscillation of the reciprocal lattice is equivalent to an oscillation of the sphere of reflection. Only the limited number of points in the reciprocal lattice which lie between the limiting positions of the sphere of reflection will reflect, and the reflection of each point will occupy only a limited fraction of the time of the oscillation. An oscillation photograph is equivalent to part of a complete rotation photograph (plates 2 (*b*), 3 (*b*).)



∘ points which contribute to powder rings.
• points outside limiting sphere, which cannot reflect.

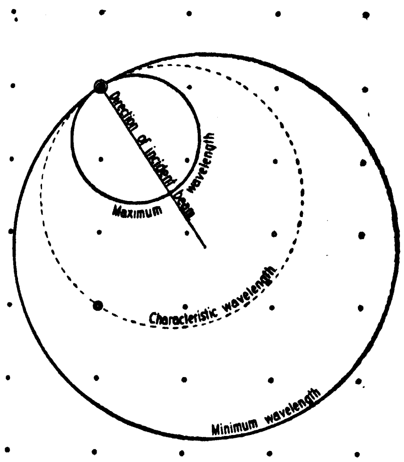
Figure 6. Powder method.

(3) A crystal *powder* may be used, the separate crystallites being uniformly oriented in all directions. Each reciprocal-lattice point becomes a sphere of radius n/d whose centre is the origin, and all such spheres intersect the sphere of reflection in a set of concentric circles. The typical powder photograph is obtained by projecting these circles on to a cylindrical film (usually a narrow strip) or on to a plane film which is normal to the incident beam. If the film is not normal to the incident beam, the powder curves become ellipses or even open conics (Lonsdale and Smith, 1941 a).

Method for use with white radiation (*Laue method*)

(4) The crystal itself may be kept stationary provided that a continuous range of X-ray wave-lengths is used. This range may, but need not, include those wave-lengths characteristic of the target. The minimum wave-length from an X-ray tube depends on the applied voltage, but is usually of the order of 0·35 A. (corresponding to a large sphere of reflection), while the maximum will be, perhaps, more than ten times as long (too long for reflection from any crystal plane). The intensity of the continuous radiation is by no means uniform. Quite apart from the possible excitation of characteristic radiation, the continuous-radiation curve shows a very broad hump of maximum intensity in the short wave-length region. All reciprocal-lattice points lying between the spheres of reflection of maximum and minimum radius will reflect during the whole time of exposure, but the intensities of the reflections will depend upon the wave-length utilized as well as upon the intrinsic reflecting power of the lattice point. Any reflections due to characteristic radiation will be of outstanding intensity.

If the incident beam lies along a symmetry axis or in a symmetry plane of the crystal, a Laue photograph taken on a plate or film placed perpendicular to the beam will reveal the appropriate crystal symmetry, and some beautiful patterns may be obtained in this way.



reflecting points •   • non-reflecting points.
strong reflection *   ◉ origin of reciprocal lattice.

Figure 7. Laue method.

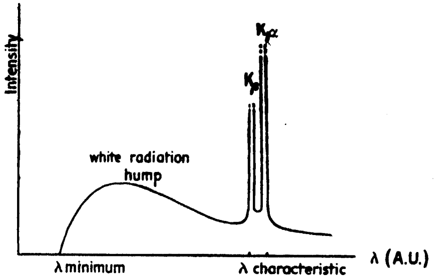


Figure 8.

It is not, of course, necessary to employ a photographic method of recording the reflections. Such methods have the advantage of giving a rapid survey and a permanent record, in a small space, of a large number of reflections, but the ionization spectrometer method may be a more accurate way of measuring exact intensity relationships.

Application of reciprocal-lattice methods to finite and irregular crystals

Since it is clear that in practice a crystal is neither infinite in extent nor perfectly regular in structure (the atomic oscillations alone must partially destroy such regularity), the question may be raised as to whether it is still justifiable to use the reciprocal-lattice method to represent finite or irregular sets of planes. This question receives an affirmative answer from the work of Laue (1936), Ewald (1940) and other mathematical physicists. The reciprocal method may be applied to any density distribution whatever, but unless there are infinite, regular sets of planes in existence, there cannot be a unique lattice of reflecting *points* in reciprocal space. The *limiting* of the planes in any direction in real space will lead to an *extension* of reflecting power of each point, along that direction in reciprocal space. Similarly, a disturbance of the regularity of the planes in real space, whatever form it may take, will cause a reduction of the reflecting power at the reciprocal-lattice points and the appearance of new reflecting points in their neighbourhood. Such a spreading of reflecting power in reciprocal space has been observed by various methods and has received a variety of theoretical explanations, all of which involve a limitation of size or of regularity of the crystal lattice.

§ 3. HISTORICAL ACCOUNT OF EARLY EXPERIMENTAL OBSERVATIONS

On Laue photographs of KCl taken by Friedrich in 1913, there appeared radial streaks which were not part of the normal diffraction pattern and which were not explained at the time. In 1922 and 1923, Clark and Duane discovered anomalous reflections from KI and other crystals, which they called x peaks, and which they explained as being due to secondary emission of *x rays characteristic of atoms* (such as iodine) *in the crystal itself*. A critical examination of their experimental method shows that the x peaks were almost certainly *extra reflections* given by the white radiation hump in the primary beam. Attempts to observe the x peaks photographically (Wyckoff, 1923) showed only smeared-out and indefinite maxima. In 1923 Faxén suggested that the Friedrich streaks might be due to the x-ray interference effects of atoms displaced by heat motion from their normal positions, and his theory, which was applied to cubic crystals only, was generalized by Waller (1925), and has become known as the Faxén-Waller theory. Waller used semi-classical arguments ; his theory was confirmed by Laue (1926), who pointed out in addition that the secondary radiation scattered by a moving atom would have suffered a change of frequency. In 1935 the theory was re-examined by Ott, using quantum methods, and he found that the formulae given by Waller were correct, to the degree of approximation adopted in the treatment. The Faxén-Waller theory differed from the Debye theory in predicting that the heat movements of the atoms would give intensity maxima, weak and diffuse by comparison with the Bragg reflections, but definitely dependent upon crystal orientation. In 1938 Laval, who saw the experimental implications of the Faxén-Waller theory perhaps more clearly than did the mathematicians themselves, planned and successfully carried out an ambitious scheme of research to determine whether the Debye or the Faxén-Waller theory most nearly corresponded to the experimental facts. He realized that the extra intensity maxima, if present at all, would persist even when the crystal was not in the proper reflecting position for a Bragg reflection, and he therefore looked for them using an ionization spectrometer method, monochromatizing his radiation (CuKα and MoKα) and arranging to be able to turn his crystal specimens into any desired orientation. Laval, whose work was carried out in Prof. Mauguin's laboratory, published a preliminary account of the results in the *Comptes rendus* in 1938, and a full account in the *Bull. Soc. Franç. Min.* in early 1939. Unfortunately the latter journal is not an easily accessible one to most physicists; his work was therefore overlooked for some time, and has not yet received the attention it deserves (see, however, Jauncey and Baltzer, 1941; W. L. Bragg, 1941; Lonsdale and Smith, 1941). It is all the more important to emphasize that Laval discovered nearly all the major experimental facts concerning the extra reflections, and that he did so as the result of a well planned scheme of research. He found (1) that discrete intensity maxima as predicted by the Faxén-Waller theory do actually exist; (2) that they can be observed by using monochromatic (characteristic) radiation with the crystal in positions not far from the Bragg setting for any given plane; (3) that their intensity (of which he obtained absolute measurements) is in general much less than that of the Bragg reflections, but increases considerably with rise in temperature; (4) that the intensity rapidly diminishes as the crystal is moved away from the Bragg setting (he gave various intensity curves showing this effect), but (5) that in the Bragg setting the extra reflection forms a shoulder on all sides of the Bragg reflection (he used a powdered specimen of silver to illustrate the broadening effect on the powder reflections). (6) He pointed out that this extra reflection corresponded to a spreading of reflecting power in reciprocal space around each reciprocal lattice point, and (7) by measuring the contours of the "isodiffusion surfaces" (surfaces of equi-reflecting power) about some Bragg reflections he showed that these surfaces are not spherical. (8) He showed that in some cases there was an extension of reflecting power right across from one reciprocal-lattice point to another, and (9) that the intensity of extra reflection for high orders was greater than it should be if extra and Bragg reflections obeyed the same law of intensity decay. (He found the following intensity values, referred in each case to a maximum of 100, for the first five observable orders of the KCl cube plane:—

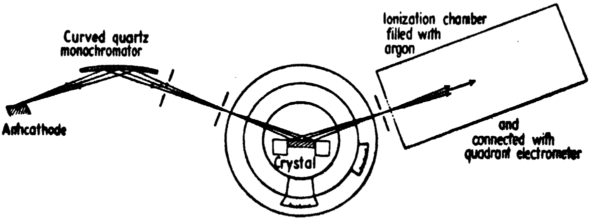

Curved quartz monochromator
Anticathode
Crystal
Ionization chamber filled with argon
and connected with quadrant electrometer

Figure 9. Laval's apparatus (*Bull. Soc. Franç. Min.* **62**, 139 (1939)).

Bragg reflection: 100 : 20 : 7 : 3 : 1
Extra reflection: 100 : 50 : 33 : 21 : 12

at a constant angle of deviation from the appropriate Bragg setting for each extra reflection.) (10) He observed that, in general, the extra reflections varied in sharpness and intensity from one substance to another, even for planes of equal intrinsic reflecting power; and finally (11) he measured the weak background intensity which is due to incoherently scattered radiation of the Compton type and also to extra reflection of wave-lengths other than those characteristic of the target. (Such unwanted wave-lengths are present even in a "monochromatized" beam, sometimes to the extent of 1 per cent or more: cf. Robinson (1933).) Laval did not apparently observe the geometrical shifts which are a feature of the extra reflections from some crystals, and he examined only elements and ionic compounds, but there is no doubt that had his results been easily available to physicists at the time of their publication, they would have aroused considerable interest and would have prevented subsequent claims for priority which have unfortunately been made.

Independently, and only a few months later, Preston (1939), while investigating the age-hardening of Cu—Al alloys by x-ray methods, observed on Laue photographs of pure Al, taken with radiation from a silver target, some weak spots which were not part of the normal diffraction pattern from such a face-centred cubic crystal. The intensity of these extra spots was enormously enhanced by raising the temperature of the specimen to 500° c.; and when the radiation was monochromatized, the extra spots remained, although all Laue spots disappeared. He found similar temperature-sensitive extra spots on Laue photographs taken in different settings, using NaCl and MgO as well as the original Al crystal. He believed that the extra spots were due to the heat motions of the atoms, but suggested that a simple approach to the theory of the phenomenon might be made by supposing the crystal to be broken up by the heat waves into small groups, each containing a few atoms all scattering in phase, each group being slightly out of phase with neighbouring groups. This method has been developed into the "simple diffraction method", which will be considered in a later section.

In 1938, Wadland, in America, had also found unexpected streaks and spots on Laue photographs of NaCl and KCl taken with tungsten radiation. He offered no explanation of these extra reflections, but Zachariasen (1938) suggested that they might be due to two-dimensional diffraction effects in the surface layer of the crystals. Late in 1939, however, at a meeting of the American Physical Society, Zachariasen withdrew this suggestion in favour of the Faxén-Waller theory, of which he offered a simplified, but admittedly less precise, derivation. The formulae he obtained included an estimate not only of the position, but of the half-width and intensity variation with increasing deviation from the Bragg setting. These formulae were compared with experiment, but although agreement was obtained in some test cases (Zachariasen, 1940; Siegel and Zachariasen, 1940; Siegel, 1941), the fact remained that since his formulae neglected the difference of velocities of elastic waves travelling in various directions in the crystal, they could, at best, only be expected to give a very rough approximation to the truth. He has since improved them by introducing the velocities and polarizations of the waves (Zachariasen, 1941 b; Jauncey, Baltzer and Miller, 1941).

Early in 1940, Sir C. V. Raman and P. Nilakantan, in an attempt to determine whether there is any x-ray effect analogous to the Raman effect in light-scattering from crystals, discovered extra reflections on Laue photographs of diamond, which they published as evidence of a supposedly new phenomenon. These extra reflections were rather sharper than those found by Preston for Al, NaCl, etc., but were similar in that each was associated with a normal Bragg or Laue reflection and was given by the *characteristic* radiation from the target used, Raman supposed that the incident x rays excited a characteristic (Einstein, infra-red) vibration of the crystal in all the unit cells simultaneously, but with slightly varying phase from one unit cell to the next, resulting in a coherent reflection of x rays of changed frequency by the "phase-waves" in the crystal. This theory, which will be considered in more detail later, has become known as the Raman theory. Raman and Nilakantan (1940, 1941) also published Laue photographs of sodium nitrate, rocksalt and calcite, and confirmed many of Laval's experimental conclusions, of which, however, they were unfortunately quite unaware. One important result which they established experimentally is that the temperature changes of intensity of the extra reflections and of the Bragg reflections not only vary considerably from one substance to another, but may even vary from one plane to another of the same crystal.

Laue photographs of organic compounds, if well exposed, sometimes show exceedingly strong extra reflections, and these have been studied by Lonsdale and Smith (1941), who have shown that they are similar in every respect to those given by ionic compounds and by metals and other elements. The photographs which illustrate this lecture show various features of the phenomenon, which may be summarized briefly as follows:—

(1) The symmetry of the Laue pattern, if any, is reproduced in the pattern of extra reflections ; but the "extra" pattern retains a measure of symmetry even when that of the Laue pattern has been completely removed by a small rotation of the crystal.

(2) The use of monochromatized radiation, which removes the Laue reflections and a good deal of the general background, leaves the extra reflections unaffected and considerably improves the conditions under which they can be seen and studied. It is noticeable that on a monochromatic photograph, however well exposed, little or no diffuse halo appears round the central spot, nor are there, in general, any radial streaks. The difficulty in obtaining good monochromatic photographs is almost entirely a question of exposure (time × intensity) (Gregg and Gingrich, 1941).

(3) The positions of the extra spots on a Laue photograph correspond closely (though not necessarily exactly) to the positions of the Bragg spots on an oscillation photograph taken with the Laue position as centre of oscillation. By a comparison of exposure times (remembering that the extra reflection is being recorded for the whole time of the Laue photograph, but that the Bragg reflection only occupies a fraction of the period of oscillation), it may be seen that the extra reflection is, in fact, very much less intense than the corresponding Bragg reflection, usually of the order of 0·01 to 0·0002 of the latter.

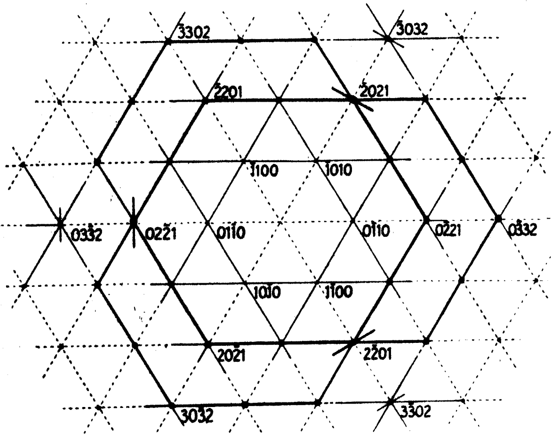
(4) The intensity of the extra reflection is greatest in the Bragg setting for a particular plane (it is then visible on a well exposed photograph only as a halo or background extension of the Bragg spot) and decreases rapidly as the angle of deviation varies in either direction. An extra reflection can sometimes be observed on successive photographs even when the crystal has been turned (say 1° at a time) through as much as 30° altogether. Such large reflecting regions are especially typical of "layer" planes (plates 4 (*a*), 5 (*b*)).

(5) A reduction of temperature to about $-180°$ c. causes a disappearance or considerable weakening in intensity of most of the diffuse extra reflections (plates 5 and 7) ; this does not apply to the sharp extra reflections which are a feature of normal (type I) diamonds, and which are also sometimes associated with strained-surface phenomena. These will be considered separately later. All temperature changes are reversible.

(6) The resemblance between the monochromatic Cu Kα Laue photograph of benzil, taken with the incident beam along the principal axis (plate 1 (*b*)) and the reciprocal lattice viewed along the same direction (figure 10) is so striking that the propriety of using the reciprocal-lattice method of studying these reflections can hardly be questioned. The comparison shows that the extra reflections in the form of streaks or spots correspond to the intersections of certain *planes* and *axes* of the reciprocal lattice with the sphere of

reflection. There are no streaks corresponding to planes passing through the origin. Most of the observed streaks extend continuously from one spot to the next, thus confirming Laval's observation that the reflecting power can extend in some cases right from one reciprocal-lattice point to the next.

In general, the main experimental facts established by the earlier work point to a weak spreading of reflecting power around each reciprocal-lattice point, sometimes coupled with a differential extension along certain reciprocal-lattice planes and axes. This extra reflecting power is markedly affected by temperature in a reversible way, but that fact alone would not justify our concluding that the effect is simply due to the temperature movements of the atoms. It might possibly be due to some structural arrangement or to some physical property of the crystals which is changed or affected by those heat movements.



_____ reciprocal-lattice planes which correspond to " extra " streaks.
.... reciprocal-lattice planes which do not correspond to " extra " streaks.

Figure 10. Incident beam along alternating hexagonal axis of benzil single crystal (normal to plane of paper). Compare plate 1 (*b*).

§ 4. SIMPLE DIFFRACTION THEORY

It is a well known fact that if the particles of which a crystal powder is composed are exceedingly small (10^{-6} cm. or less in average diameter) the x-ray powder reflections are broadened. Indeed, this method is used as a means of measuring the size of colloidal particles. Laue and Riewe (1936) have discussed the diffraction effects to be expected from finite crystalline particles, and their results were extended and presented in a new way by Ewald at the Physical Society's 1939 (Bristol) Conference on the Solid State. For a crystal particle of any limited shape, the reflecting power in reciprocal space is not confined to sharp points, but is extended along directions normal to the boundary planes of the particle. Thus the regions in reciprocal space corresponding to a crystal of needle shape are discs normal to the length of the needle, and conversely a disc-shaped crystal will have needle-like reflecting regions instead of points in reciprocal space. As long as the particle is not so small that the parameters of the internal structure are comparable with the dimensions of the external form, a spherical particle will have spherical regions of reflecting power in reciprocal space.

One most important point to notice is that *all the reflecting regions in reciprocal space corresponding to a finite crystal particle* (or to an assemblage of finite, parallel, but independent crystal particles) *will be similar and similarly oriented.*

If a large, apparently single, crystal is actually composed of very small parallel crystallites, the reflecting regions in reciprocal space, besides being similar and similarly oriented, will be entirely diffuse. There will be no sharply defined reflecting regions such as are indicated by the sharp Laue spots. If the crystal is made up of a mixture of large and very small crystallites there will be sharp reflecting regions, each with an extended diffuse background.

The result of supposing, as Preston did, that the crystal diffracts not as a whole but as a system of small blocks all slightly disoriented with respect to the average crystal directions or having slightly different spacings, is to postulate the existence of a number of reciprocal lattices of slightly differing sizes and orientations, all having entirely diffuse reflecting regions. Near to the common origin the superposed diffuse reflecting regions will be rather similar to those of a system of parallel small crystallites of uniform spacing, but regions farther from the origin will be extremely diffuse and extended. This is not at all in agreement with experiment.

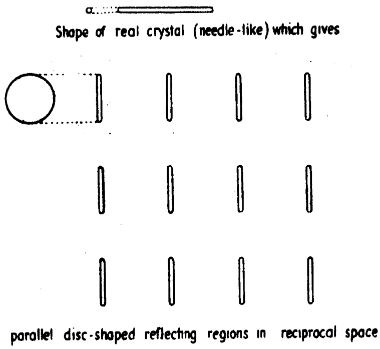

Shape of real crystal (needle-like) which gives
parallel disc-shaped reflecting regions in reciprocal space

Figure 11.

The method has been developed by Sir W. H. Bragg (1940, 1941), who has shown that a great many of the observed phenomena can be geometrically accounted for by postulating the existence of two-dimensional sheets of atoms in the crystal, all x-ray scattering from the atoms in any one sheet being in phase, but with a random phase-relationship between such sheets and the remainder of the crystal. The reciprocal-lattice method was not used by Sir W. H. Bragg, who calculated the diffraction effects directly, but it may readily be seen that such sheets in real space correspond to intensity spikes in perpendicular directions in reciprocal space and that, by symmetry, any one such spike parallel, say, to the [110] direction of a cubic crystal must be accompanied by spikes along all crystallographically equivalent directions, for each reciprocal-lattice point.

Now it will be shown that the experimental data do, in fact, correspond either to ellipsoidal regions having their main axes along certain main crystallographic directions, or to intensity spikes along certain of these directions in reciprocal space; and up to a point, therefore, it will certainly be possible, by choosing suitable groups or arrangements of atoms scattering in phase with each other, but out of phase with their neighbours, to explain observed experimental effects. It is found, however, that in explaining those extra spots that *are* observed, the theory also predicts some extra spots of equal or greater intensity, which are definitely not present. Nor is it possible, by means of the theory, to predict the detail of the extra reflections for unknown crystals even of the simplest composition or structural arrangement. As it stands, the simple diffraction theory does not explain why all body-centred cubic metals do not behave similarly, nor why face-centred cubic metals may or may not behave like body-centred cubic metals. Nor does it appear to offer any adequate explanation of the large temperature changes of intensity which occur.

Nevertheless, it will be seen later that this theory, although inadequate to explain all the effects, has a certain significance; and some form of the theory probably provides the correct explanation of those temperature-independent, structure-sensitive extra spots which appear when the crystal investigated has a freshly cleaved or worked surface, or is subject to certain types of strain.

§ 5. FAXÉN-WALLER (THERMAL) THEORY

In the elastic-vibration, or thermal theory, first correctly developed along semi-classical lines by Faxén and Waller, use was made of the known fact that the atoms of any crystalline substance are continuously in motion about their mean positions. It is not possible, within the scope of a survey of this kind, to give a rigorous account of what is really a most complicated mathematical treatment, but the salient points of the argument, which has been further elaborated and given a strictly quantum-mechanical basis by Laue (1926), Ott (1935), Zachariasen (1941), Born and Sarginson (1941, 1942), Jahn (1942) and others, will here be given as simply as possible.

The principal assumption made is that the forces in the crystal are harmonic. The frequency of the atomic vibrations (say 10^{13} sec.$^{-1}$ or less) is very low compared with that of the x rays (about 10^{18} sec.$^{-1}$). To the incident x rays, therefore, the crystal will appear to be stationary, but with the atoms displaced unequally from their mean positions. These atomic displacements can be resolved into a system of standing waves superimposed upon the regular crystalline structure. The time-integrated effect of any one wave can then be considered separately, and the results of all the waves added together.

The structure factor (and hence the intensity of reflecting power, which is proportional to the square of the structure factor) of any set of parallel planes (*hkl*) depends upon the relative positions of the atoms in the direction normal to the planes. The structure factor will be changed if the atoms change their positions in respect of their relative distances from the planes, but not if they merely move parallel to the planes. It may be shown that in the case of a harmonic vibration in which the atomic movements are *normal* to a given set of planes, the resulting change in the intensity of reflecting power is, to a first approximation, proportional to the *mean-square amplitude*, $\overline{\xi^2}$, of the oscillation. The mean energy of an oscillator, $\overline{E}$, is proportional to $\nu^2\overline{\xi^2}$, where ν is the frequency, and $\overline{E}$ for a Planck oscillator is

$$h\nu\left(\frac{1}{2}+\frac{1}{e^{h\nu/kT}-1}\right).$$

If $h\nu \ll kT$, $\bar{E}$ therefore approximates to kT. *The intensity change of reflecting power* due to the oscillation is therefore proportional to kT/ν^2, and *is greater the lower the frequency*.

The possible vibrations of a lattice may be roughly classed as of two kinds: the acoustical vibrations, in which neighbouring atoms move, in general, in the same direction, and the optical vibrations, in which neighbouring atoms move, in general, in opposite directions. In both cases, the frequency of the vibration is a function of its wave-length, but in the case of the acoustical vibrations the frequency becomes zero for certain values of the wave-length, while for the optical vibrations the frequency is never less than a rather high minimum (which may or may not be less than the maximum acoustical frequency *).

Since it is the lowest frequencies which correspond to the largest amplitudes, and which are, therefore, of greatest importance in assessing the intensity changes due to the atomic vibrations, and since these lower frequencies are entirely those of the acoustical spectra, it is only the acoustical (or elastic) vibrations which need be considered in the first instance. A complete theory of the effect must, however, take account of the whole frequency spectrum, including both acoustical and optical branches, which are not always distinguishable from each other.

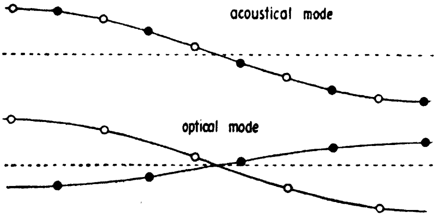


Figure 12.

The frequency of a single wave is equal to its velocity divided by its wave-length. The velocity of an elastic wave of given wave-length, travelling in any given crystal direction, is determined only by the wave-length, the density and elastic constants of the crystal, and the polarization of the wave. The wave-length can, in practice, vary from infinity down to a length comparable with the unit-cell size. Along certain main crystal directions, the elastic waves can be resolved into three components, two transverse and one longitudinal; but, in general, such a resolution is not possible.

Let us, to take a simple case, consider the effect of a transverse stationary vibration whose wave-length μ is along the [100] direction of a cubic crystal (figure 13). Clearly there is no atom along [100] which exactly repeats the conditions at A until we come to A′, which is one wave-length away. The effect of the stationary wave is to impose a super-lattice on the original structure, the dimensions of the super-lattice being dependent upon the wave-length and direction of the stationary wave (in figure 13, μ is assumed, for the sake of convenience, to be an integral multiple of the lattice spacing, but this is not necessary).

---

* For a more precise account of lattice vibrations see Born and Kármán (1912), Born and Göppert-Mayer (1933) or the paper by Professor Born on p. 362 of this issue.

In the super-lattice, though not in the original lattice, there will be planes AB′, AC′ which are nearly of the same orientation as the plane AA′, and *also nearly of the same spacing*. The longer the wave-length μ, the nearer is this approximation. To a first approximation, only the one pair of planes, AB′ and AC′, in the super-lattice, will give reflections of appreciable intensity, and thus the result of the stationary wave is to introduce a pair of " ghost " reflections, one on each side of the original reflection from AA′. In reciprocal space there will be, on each side of the lattice point P corresponding to the set of planes AA′, a weakly reflecting point whose distance from P is $1/\mu$; in fact, this will be true for every lattice point, but the intensity of the ghosts will vary with the points.

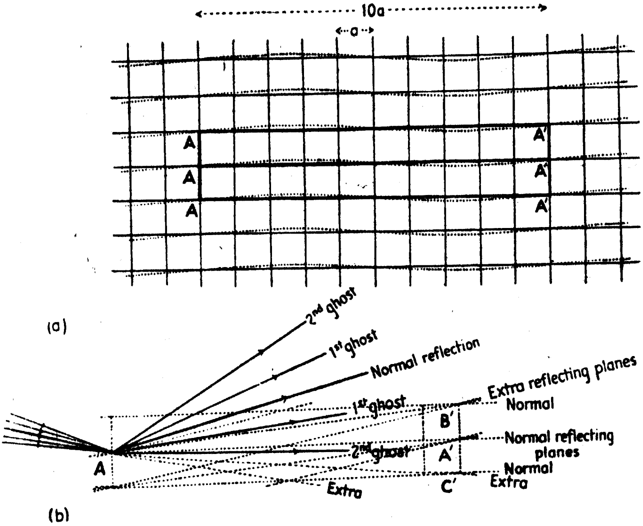


Figure 13.

(*a*) Showing super-lattice due to presence of transverse wave of length 10 *a*.

(*b*) Showing extra reflections from new planes introduced by super-lattice (angles exaggerated).

Now there are, of course, many such stationary waves to be considered, and these have various directions, wave-lengths and polarizations. Each imposes on the crystal a super-lattice and each atom in the crystal belongs simultaneously to all these super-lattices. Each super-lattice introduces a pair of extra reflecting points for every reciprocal-lattice point, and hence *each reciprocal-lattice point* (taking all the super-lattices into account) *will be surrounded by a cloud of weak extra reflecting power*, its own reflecting power being correspondingly reduced. Will the distribution of intensity of this cloud be spherical for all lattice points ? Only if the velocity (which is determined by the frequency) of an elastic wave of given wave-length and polarization is the same whatever its direction of propagation in the crystal. This condition is never fulfilled, even approximately,* as will be shown later.

Waves, whether longitudinal or transverse (or a mixture of both), travelling along a given crystal direction in real space, give extra reflecting points *along that same direction* for each lattice point in reciprocal space, but the intensity of reflecting power along these extended directions depends not only on the frequencies but also upon the polarizations of the waves, that is, on the direction of the atomic movements relative to the various crystal planes. This introduces differences between different reciprocal-lattice points which will be described in more detail later.

---

* The assumption of such a condition has wrongly led Raman to reject the thermal theory as being inconsistent with experiment. Theorists have often, unfortunately, assumed this condition for simplification of the calculations.

The intensity of reflecting power at any point in the neighbourhood of a reciprocal-lattice point is also dependent upon the total angle of deflection 2θ. This angle enters into the intensity expression in a number of ways (see Born and Sarginson, 1941), some of which involve a *decrease* and others an *increase* of intensity with increasing 2θ. For instance, the polarization factor $(1+\cos^2 2\theta)/2$ and the atom-scattering factor f (which is dependent upon $\sin\theta/\lambda$), which are common both to Bragg and to extra reflections, involve a decrease of intensity with increasing 2θ. So does the modification of the atom-scattering factor due to the atomic vibrations, a factor which changes f into fe^{-M} where M is dependent upon $\sin\theta/\lambda$, absolute temperature T and the interatomic forces (Debye, 1914; Waller, 1925; Waller and James, 1928; James, Waller and Hartree, 1928). On the other hand, a given atomic oscillation will cause more spreading of reflecting power for planes of small than for those of large spacing, and this fact enters the intensity expression as a term $(1-\cos 2\theta)=2\sin^2\theta\sim\frac{1}{2}(n\lambda/d)^2$, where d is the spacing corresponding to the nearest lattice point.

The X-ray wave-length λ only influences the intensity indirectly, through θ and through the term $\sin\theta/\lambda$ which enters into the expression for the atom-scattering factor and its temperature modification.

The extra intensity is also very dependent upon the temperature T. As mentioned before, the amplitudes of the oscillations are influenced by temperature because of the increase of mean energy $\boldsymbol{h}\nu\left(\frac{1}{2}+\frac{1}{e^{\boldsymbol{h}\nu/\boldsymbol{k}T}-1}\right)$ with T (for vibrations of low frequency the mean energy is $\boldsymbol{k}T$), and the intensity is controlled by the mean-square amplitude in two ways: (1) directly, as a result of the actual atomic displacements; (2) indirectly, as an exponential modification of the atom-scattering factor (see last paragraph). Before the law of dependence upon temperature can be experimentally tested, however, it is necessary to know what range of frequencies is operative in giving any particular extra spot, and, therefore, we must consider in rather more detail the shape of the contours of equal reflecting power around any particular lattice point, and the dependence of extra intensity in the neighbourhood of the lattice point upon the nature of the crystal.

In order to do this we shall consider the simple case of a harmonic wave travelling along a main crystallographic direction, and determine the effect on the reflecting power in the neighbourhood of various reciprocal-lattice points. Leaving the variation with 2θ and T out of account, we know that

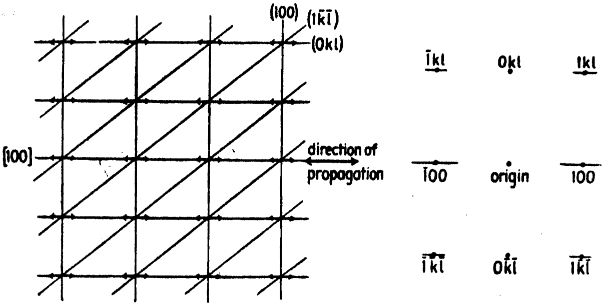
(1) Elastic waves propagated along a given direction in the crystal cause an extension of reflecting power of each reciprocal-lattice point along the direction of propagation. Each wave-length μ contributes to the reflecting

power at a definite point along this direction (or rather a pair of points, one on each side of the lattice point, which itself corresponds to waves for which $1/\mu = 0, 1/a, 2/a \ldots$).

(2) The intensity of reflecting power at any given point depends on $1/(\text{frequency})^2$ of the contributing waves. This factor introduces both the wave-length μ and the elasticity of the crystal in respect of the atomic movement involved.

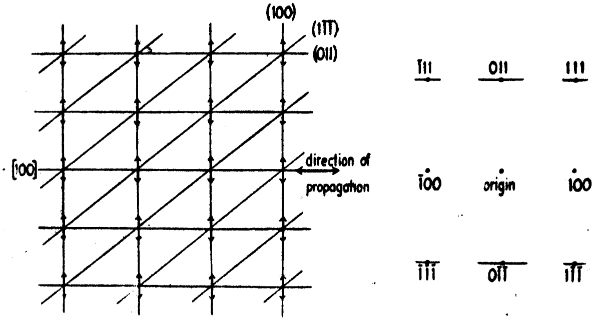
(3) The intensity also depends upon the square of the cosine of the angle between the radius vector of the given point in reciprocal space and the direction of the atomic movements (that is, upon the polarization of the waves).

In figure 14 there is shown schematically a longitudinal wave travelling along the [100] direction of a cubic crystal. The extension of reflecting power for *each* reciprocal-lattice point will be along [100]. Since the atomic movements take place along [100], that is, *parallel* to the ($0kl$) planes, the intensity of extra reflecting power for all ($0kl$) planes will be zero, whereas that for the (100) planes will be a maximum (since the atoms move *perpendicular* to the (100) planes).



(*a*) Real lattice, longitudinal vibration.
(*b*) Reciprocal lattice, extension of reflections.

Figure 14.



(*a*) Real lattice, transverse vibration.
(*b*) Reciprocal lattice, extension of reflections.

Figure 15.

For general planes (*hkl*), the intensity will be intermediate in value. Since the atomic movements only involve simple compression or expansion along a cube direction, the velocity of the wave is dependent upon c_{11}, in fact it equals

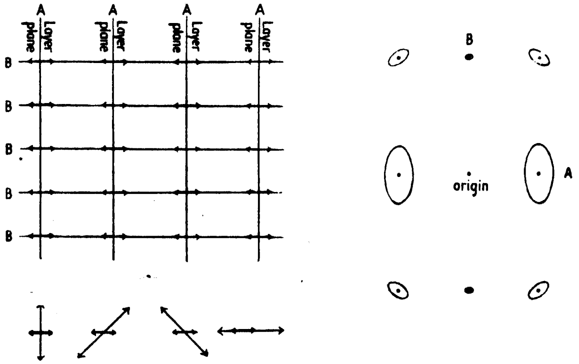
$\sqrt{(c_{11}/\rho)}.f(\mu)$, where ρ is the crystal density and $f(\mu)$ a function of the wave-length.

A transverse wave travelling along the same [100] direction (figure 15) will have quite a different effect on the reflecting power of the various planes. The atomic movements are now *parallel* to the (100) planes, and the reflection from those planes is therefore not affected. They are *perpendicular* to one of the (0***kl***) planes, say (011), and the reflecting power of (011) is therefore extended along the [100] direction. The atomic movements involve simple shear of the (100) planes, and the velocity of such a wave will be $\sqrt{(c_{44}/\rho)}.f(\mu)$. Since c_{44} is always less, and may be very much less, than c_{11}, the velocity of transverse waves is always less than that of longitudinal waves of equal length, along the cube directions of a cubic crystal.

By such simple considerations as these, it is possible to arrive at general conclusions concerning the extra reflections from "layer" and "chain" structures.

Layer structures

The atoms in the layer planes A of a structure such as graphite, hexamethylbenzene, etc., are more densely packed and more closely bound than in other crystal planes. Such planes are often cleavage planes, the forces between them are weak van der Waals forces, and can be easily disrupted to permit gliding, or even the introduction of foreign atoms. Layer structures are most compressible in a direction normal to the layer plane and bend most easily about directions in that plane. In other words, the atoms move most easily in the direction of the normal to the layer plane. Waves polarized in this way will have the biggest amplitudes no matter what their directions of propagation, and, therefore, *the layer plane will show a maximum extension of reflecting power in every direction in reciprocal space*, and a maximum temperature change. The favoured direction of atomic movement is parallel to planes, such as B, which are normal to the layer plane, and such planes, therefore, will show a minimum extension of reflecting power in all directions.



Preferred modes of vibration have amplitudes normal to layer plane for all directions of propagation.

Reflecting regions in reciprocal lattice of a LAYER structure.

Figure 16.

This is fully borne out by observation. Laue photographs of layer structures are characterized by one very large, intense and persistent extra reflection, which must correspond to a huge spreading of reflecting power about the reciprocal-lattice point of the layer planes. Graphite, for instance, shows an enormous extra reflection for all observable orders from its cleavage plane, and these completely disappear at liquid-air temperatures, while all other planes give quite small extra reflections (plates 4 (*a*), 5 (*a*) and (*b*)).

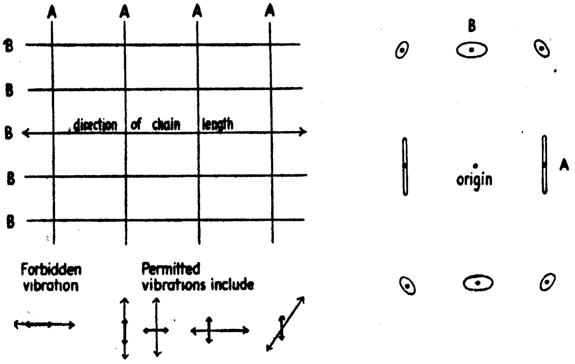
The "background" of the Bragg reflection from a layer plane will sometimes appear even on a rotation or Weissenberg photograph, although as a rule the exposure given to such a photograph is insufficient to show the background effects.

It is noteworthy that Hughes (1941), in determining the structure of melamine by means of a Fourier analysis, found it necessary to apply a differential temperature correction for structure factors of planes parallel or normal to the molecular planes of this layer structure. The spreading of reflecting power for planes nearly parallel to the molecular layers was sufficient to reduce the observed structure factors quite materially even at room temperatures. Raman and his colleagues (1941 Symposium in the *Proc. Ind. Acad. Sci.*) have observed that the extra reflections can be quite strong even for planes which have a relatively weak intrinsic reflecting power, provided that those planes are nearly parallel to the molecular layers. They instance this as being a proof that the optical vibrations are responsible for the extra reflections ; but it is, in fact, a consequence of the elastic vibration theory, since the amplitudes depend on the quasi-elastic interatomic forces, not on the static structure factor of the plane under consideration, and these forces are very anisotropic both for layer structures and for the chain structures which will be next considered. The misunderstanding has, of course, arisen from Raman's incorrect assumption that the elastic forces are equal in all directions for waves of all kinds.

The existence of a differential temperature effect means also that the relative intensities of Bragg and extra reflections may vary considerably with temperature for different crystal planes, as has been found (Kracek, Posnjak and Hendricks, 1931 ; Raman and Nilakantan, 1940).

Chain structures

Müller (1941) has shown that for long-chain compounds the compressibility along the length of the chain is several times smaller than that in directions at right angles. This means that longitudinal vibrations along the chain axis cannot take place. This prohibition does not affect the spreading of reflecting power about the points B, which will take place in all directions (the forbidden movements would have been parallel to the planes B) ; but it means that there can be *no extension of reflecting power for points A along the chain direction.* The surfaces of equal reflecting power for A and for lattice points corresponding to planes nearly normal to the chain direction will be *discs*, while those for B and neighbouring lattice points will be more nearly spheroidal.



Reflecting regions in reciprocal lattice of a CHAIN structure.

Figure 17.

The extra spots corresponding to planes such as B will be large and more nearly circular (figure 18), while those corresponding to planes such as A will be fine lines perpendicular to the chain direction (figure 19). Sometimes it is possible to observe both types of spot on one Laue photograph (plate 4 (*b*)). Such photographs as these not only allow us to detect at once the presence of a chain structure, but show which planes are normal to the chain direction and so form a useful subsidiary method of structure determination. In structures which consist of chains of lath-like molecules, the orientation of the molecules (although not their relative translations) can be settled entirely by observation of the extra reflections (Lonsdale, Robertson and Woodward 1941).

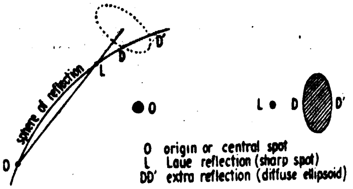

sphere of reflection
O
L
D
D'
O
L
D
D'
O origin or central spot
L Laue reflection (sharp spot)
DD' extra reflection (diffuse ellipsoid)

Figure 18.

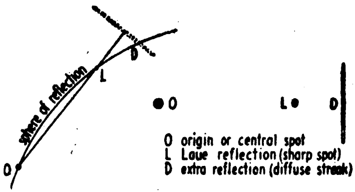

sphere of reflection
O
L
D
O
L
D
O origin or central spot
L Laue reflection (sharp spot)
D extra reflection (diffuse streak)

Figure 19.

Cubic crystals

Although the general Waller formula is rather unmanageable, it can be expressed in a relatively simple way for cubic crystals of anisotropic elastic properties; this has been done by Zachariasen (1941) for some special cases and independently by Jahn (1942), whose general formula is now given * :

$$
\begin{aligned}
I \propto \frac{R^2}{r^2}\, \{c_{44}^2 &+ L^2[c_{44}(c_{11}-c_{44})(m^2+n^2)+(c_{11}+c_{12})(c_{11}-c_{12}-2c_{44})m^2n^2] \\
&+ M^2[c_{44}(c_{11}-c_{44})(n^2+l^2)+(c_{11}+c_{12})(c_{11}-c_{12}-2c_{44})n^2l^2] \\
&+ N^2[c_{44}(c_{11}-c_{44})(l^2+m^2)+(c_{11}+c_{12})(c_{11}-c_{12}-2c_{44})l^2m^2] \\
&- 2MNmn(c_{12}+c_{44})[c_{44}+(c_{11}-c_{12}-2c_{44})l^2] \\
&- 2NLnl(c_{12}+c_{44})[c_{44}+(c_{11}-c_{12}-2c_{44})m^2] \\
&- 2LMlm(c_{12}+c_{44})[c_{44}+(c_{11}-c_{12}-2c_{44})n^2]\} \div \\
\{c_{11}c_{44}^2 &+ c_{44}(c_{11}+c_{12})(c_{11}-c_{12}-2c_{44})(l^2m^2+m^2n^2+n^2l^2)+ \\
&\qquad (c_{11}+2c_{12}+c_{44})(c_{11}-c_{12}-2c_{44})^2 l^2m^2n^2\}.
\end{aligned}
$$

I is the intensity of extra reflecting power at a distance r in the direction $[l, m, n]$ from a reciprocal-lattice point P, distant R from the origin of reciprocal space,

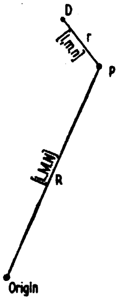


Figure 20.

the direction cosines of R being $[L, M, N]$; c_{11}, c_{12}, c_{44} are the elastic constants of a cubic crystal, in terms of which the compressibility is $(c_{11}+2c_{12})/3$ and the shear constant c_{44}. The crystal is said to be elastically isotropic if $c_{11}-c_{12}=2c_{44}$, and in that case the formula reduces to

---

* This formula applies to a cubic crystal of any lattice, simple, face-centred or body-centred, if the optical vibrations are neglected.

$$I \propto \frac{R^2}{r^2}\left\{\Sigma L^2\left[\frac{1}{c_{11}}+\left(\frac{1}{c_{44}}-\frac{1}{c_{11}}\right)(m^2+n^2)\right]-2\Sigma MNmn\left(\frac{1}{c_{44}}-\frac{1}{c_{11}}\right)\right\}.$$

Hence, for a given I and R,

$$r^2 \propto \frac{1}{c_{11}}+\left(\frac{1}{c_{44}}-\frac{1}{c_{11}}\right)[1-(Ll+Mm+Nn)^2]$$

$$\propto \frac{\cos^2\theta}{c_{11}}+\frac{\sin^2\theta}{c_{44}},$$

where θ is the angle between the vectors R and r.

It will readily be seen that, for a given value of R and I, r^2 can never be constant unless $c_{11}=-c_{12}=c_{44}$ (condition for spherical spread of reflecting power). In that case, however, $c_{11}+2c_{12}=-c_{44}$, involving a negative compressibility or a negative shear constant, either of which is absurd. The surfaces of equal reflecting power (isodiffusion surfaces) cannot be spherical. In fact, even in the unusual event of a negative c_{12}, the highest value that c_{44} can have in an isotropic cubic crystal is $3c_{11}/4$; if c_{12} is positive then $c_{44}<c_{11}/2$ (Jahn and Lonsdale, 1942).

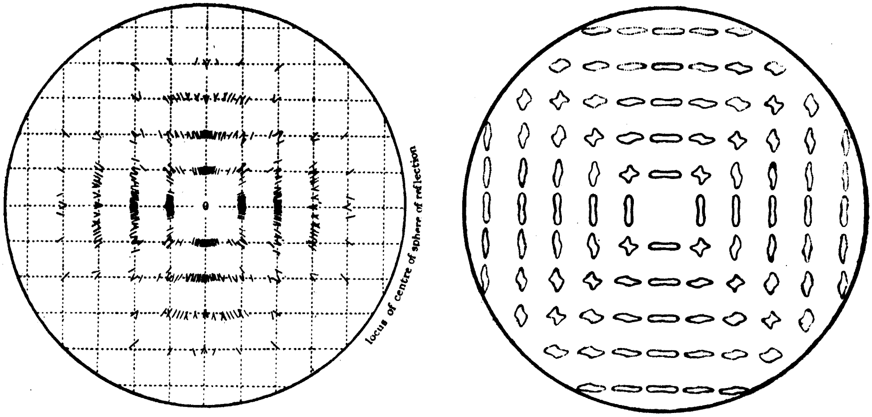
The presence of R^2 in the formula would make it appear that the intensity of extra reflection must increase rapidly as the spacing of the planes decreases. R^2 is, in fact, the term $(1-\cos 2\theta)$, which was previously mentioned as leading to an increase of intensity with increasing θ. It must be remembered that the Jahn formula given above does not introduce those factors (polarization factor, atom-scattering factor, etc.) which cause a decrease of intensity with increasing θ. The combined effect of such factors in practice is (as Laval has shown, and as may easily be observed) that the integrated extra reflections do in general fall off in intensity with increasing angle of deviation, but not so rapidly as the Bragg reflections. In the subsequent discussion, therefore, R^2 will be omitted.

Many interesting consequences of the formula have been given by Jahn, but perhaps one of the most instructive is its application to such crystals as KCl (for which $c_{11}-c_{12}>2c_{44}$) or to metals such as sodium or lead (for which $c_{11}-c_{12}<2c_{44}$). There are various ways of comparing formula and photograph, but in practice the intermediate use of the reciprocal lattice is almost indispensable. From the formula it is possible, by substituting values of c_{11}, c_{12}, c_{44}, L, M, N and l, m, n to plot the values of r corresponding to a constant I for each reciprocal lattice point (*isodiffusion surfaces*), or to plot the values of I corresponding to a constant r. This gives a distribution of reflecting power about each reciprocal-lattice point which, as mentioned before, always involves a maximum extra reflecting power along certain reciprocal-lattice *planes* or *axes*, but in which the relative extensions of reflecting power along those directions vary from one lattice point to another.

By calculating back from the photographs to the reciprocal lattice (analytically or geometrically), it is possible either

(1) to compare the experimentally observed extensions of reflecting power about each point, plotted from a complete series of photographs, with the picture given by the formula, as in figures 21 (*a*) and (*b*) ; or

(2) to compare the positions and intensities of the extra reflections on one Laue photograph at a time, with the intersections of the corresponding sphere of reflection with the theoretical (three-dimensional) diagram. Using the latter

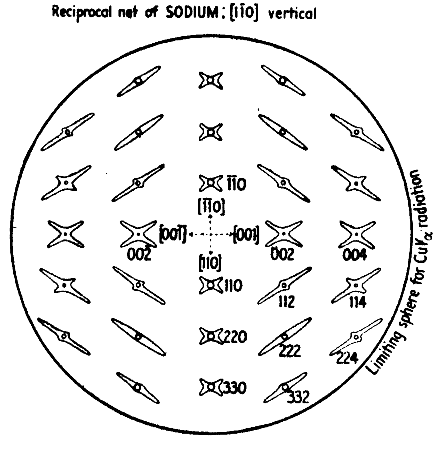


(*a*) Fig. 21. (*b*)

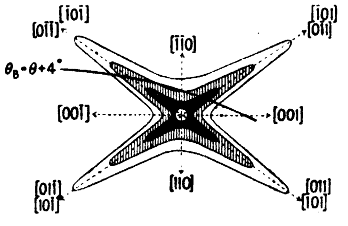
(*Reproduced by courtesy of the Council of the Royal Society.*)

(*a*) KCl ; Ag Kα : reciprocal net showing observed positions of extra spots on central layer-line of Laue photographs taken with cube axis of crystal vertical.

(*b*) KCl ; Ag Kα : reciprocal net showing values of I (r constant) for [$hk0$] lattice points, calculated from the Jahn formula. The decrease of intensity with scattering angle is indicated by decreasing line intensity.


Reciprocal net of SODIUM; [1̄10] vertical
1̄10
[1̄10]
[001̄]
[001]
002̄
002
004
[110]
110
112
114
220
222
224
330
332
Limiting sphere for CuK$_\alpha$ radiation

Showing intensity distributions (projected on (1$\bar{1}$0) plane) for various lattice points. Intensity of [1$\bar{1}$0] peak indicated by circles. Distances along other directions are proportional to I of Jahn's formula (r constant).



Enlargement of 002 reflecting region showing trace of sphere of reflection for $\theta = \theta_B + 4°$ (the four peaks shown are pointing up out of plane of paper at 30° and are accompanied by four similar peaks pointing down at 30°).

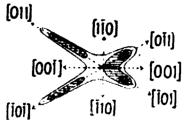


Diagram showing section of 002 reflecting region by sphere of reflection (compare plate 6 (*a*)).

Figure 22.

method, Lonsdale and Smith (1941 d, 1942 a) have shown that the observed extra reflections for Na, Pb and W single crystals correspond in detail with the predictions of the Jahn formula. To take one example (figure 22, plate 6 (*a*))

Table 1. Values of I from Jahn's formula (r constant)

$LMN\rightarrow$ lmn $\downarrow$	Na			Pb			W			KCl		
	001	$\frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}}0$	$\frac{1}{\sqrt{3}}\frac{1}{\sqrt{3}}\frac{1}{\sqrt{3}}$	001	$\frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}}0$	$\frac{1}{\sqrt{3}}\frac{1}{\sqrt{3}}\frac{1}{\sqrt{3}}$	001	$\frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}}0$	$\frac{1}{\sqrt{3}}\frac{1}{\sqrt{3}}\frac{1}{\sqrt{3}}$	001	$\frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}}0$	$\frac{1}{\sqrt{3}}\frac{1}{\sqrt{3}}\frac{1}{\sqrt{3}}$
100	24	22	22	7	5	5	0·66	0·43	0·50	15	9	11
010	24	22	22	7	5	5	0·66	0·43	0·50	15	9	11
001	19	24	22	2	7	5	0·20	0·66	0·50	3	15	11
$\frac{1}{\sqrt{2}}\ \frac{1}{\sqrt{2}}\ 0$	24	12	16	7	2	3	0·66	0·20	0·35	15	4	7
$\frac{1}{\sqrt{2}}\ \frac{\bar{1}}{\sqrt{2}}\ 0$	24	165	118	7	27	20	0·66	0·66	0·66	15	6	9
$0\ \frac{1}{\sqrt{2}}\ \frac{1}{\sqrt{2}}$	86	56	16	14	11	3	0·43	0·54	0·35	5	10	7
$0\ \frac{1}{\sqrt{2}}\ \frac{\bar{1}}{\sqrt{2}}$	86	56	118	14	11	20	0·43	0·54	0·66	5	10	9
$\frac{1}{\sqrt{2}}\ 0\ \frac{1}{\sqrt{2}}$	86	56	16	14	11	3	0·43	0·54	0·35	5	10	7
$\frac{\bar{1}}{\sqrt{2}}\ 0\ \frac{1}{\sqrt{2}}$	86	56	118	14	11	20	0·43	0·54	0·66	5	10	9
$\frac{1}{\sqrt{3}}\ \frac{1}{\sqrt{3}}\ \frac{1}{\sqrt{3}}$	30	20	10	10	6	2	0·50	0·35	0·20	6	5	4
$\frac{\bar{1}}{\sqrt{3}}\ \frac{1}{\sqrt{3}}\ \frac{1}{\sqrt{3}}$	30	41	37	10	14	12	0·50	0·66	0·61	6	8	7
$\frac{1}{\sqrt{3}}\ \frac{\bar{1}}{\sqrt{3}}\ \frac{1}{\sqrt{3}}$	30	41	37	10	14	12	0·50	0·66	0·61	6	8	7
$\frac{1}{\sqrt{3}}\ \frac{1}{\sqrt{3}}\ \frac{\bar{1}}{\sqrt{3}}$	30	20	37	10	6	12	0·50	0·35	0·61	6	5	7
(At 290° K.) $10^{11}\times$	c_{11} 0·521	c_{12} 0·400	c_{44} 0·410	c_{11} 4·77	c_{12} 4·03	c_{44} 1·44	c_{11} 50·7	c_{12} 20·2	c_{44} 15·2	c_{11} 3·88	c_{12} 0·64	c_{44} 0·65

Table 2. Observed and calculated extra intensities for sodium

(*hkl*)	±[*pqr*]	[100]	[010]	[001]	[110]	$[1\bar{1}0]$	[011]	$[01\bar{1}]$	[101]	$[\bar{1}01]$
(002)		— (24)	— (24)	— (19)	— (24)	— (24)	M.S. (86)	M.S. (86)	M.S. (86)	M.S. (86)
(110)		— (22)	— (22)	— (24)	— (12)	V.S. (165)	M. (56)	M. (56)	M. (56)	M. (56)
(112)		— (24)	— (24)	— (21)	— (20)	M. (71)	W. (30)	S. (133)	W. (30)	S. (133)
(222)		— (22)	— (22)	— (22)	— (16)	S. (118)	— (16)	S. (118)	— (16)	S. (118)

the (002) extra reflection for Na, when the $[1\bar{1}0]$ axis is vertical and the angle of incidence is θ_B+4 (25°·3 for CuKα), consists of what looks like *a pair of scissors*, but *with no hinge* of appreciable intensity. This corresponds to the intersection, by the sphere of reflection, of intensity spikes along directions $[\bar{1}0\bar{1}]$ $[0\bar{1}\bar{1}]$ (blades of scissors) and $[\bar{1}0\bar{1}]$ $[0\bar{1}1]$ (handles of scissors). There is no hinge because there is no intensity spike along $[\bar{1}\bar{1}0]$. Other photographs show that for the (002) reciprocal-lattice point there are no intensity spikes along $[1\bar{1}0]$ $[\bar{1}10]$, although for the (110) point the intensity spikes along $[1\bar{1}0]$ $[\bar{1}10]$ are particularly strong (plate 6 (*b*)). Others, again, illustrate the fact that for the (222) point there are no intensity spikes along [101] [011] [110] $[\bar{1}0\bar{1}]$ $[0\bar{1}\bar{1}]$ or $[\bar{1}\bar{1}0]$, although along the remaining six diagonal directions they are strong. Tables 1 and 2 show how figures 21 and 22 are deduced from the Jahn formula, and how the absence of certain spots and streaks (which was inexplicable in terms of the simple diffraction theory) is strictly in accordance with the formula, which predicts that the intensity spikes along those very directions shall be absent or weak.

Reciprocal net of TUNGSTEN, $[1\bar{1}0]$ vertical

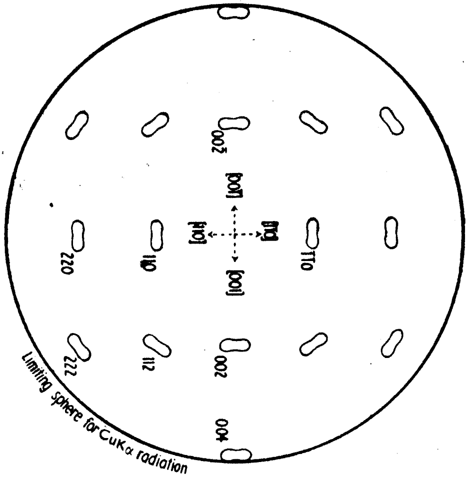


Figure 23. Relative values of *I* are shown on a scale 80 times larger than those shown in figure 22 for sodium. (Compare table 1.)

Lonsdale and Smith have found that lead (plate 7) also gives extra reflections corresponding to intensity spikes along the various [110] directions, which are similar to, although less extended than, those found for sodium, and this in spite of the fact that sodium has a body-centred cubic structure, whereas lead has a face-centred cubic structure. Table 1 shows that this similarity is predicted by the Jahn formula, which, since it neglects optical vibrations, gives similar results for cubic crystals of all classes *provided that they are elastically similar.* Sodium and lead are both elastically anisotropic in the same sense ($c_{11}-c_{12}<2c_{44}$), and their elastic constants are small. On the other hand, tungsten single crystals, although similar in crystal structure to sodium, give quite different extra reflections. Those of sodium (and lead) are intense, large and detailed in structure, and persistent over a wide range of angles of incidence; those of tungsten are weak, small, elliptical, and disappear very quickly as the angle of setting departs from the Bragg angle for each plane in turn. This is because the elastic constants for tungsten are large and isotropic ($c_{11}-c_{12}=2c_{44}$) and lead, therefore, to quite a different intensity distribution (table 1 or figure 23) around the reciprocal lattice points from that found for soft, anisotropic metals.

It is possible to see the relation of the sheets of atoms postulated in the simple diffraction theory to the actual state of affairs in the crystal when it is realized that normal to the propagation directions of the elastic waves of lowest velocity there will, indeed, be sheets of atoms scattering more nearly in phase with each other than with atoms in other directions; but the phase relationships between successive sheets cannot be ignored, since they are not by any means random. It is here that the simple diffraction theory breaks down, in that it does not introduce any physical property of the crystal which shall decide either *which sheets shall be chosen, or what are the phase-relationships between successive sheets.*

Effects due to strain, intrinsic or mechanical

It has been pointed out that, in addition to the extra reflections referred to above, which appear with equal intensity under similar conditions for any crystal of the same sort, there are sometimes extra reflections of quite another kind, which are structure-sensitive, but not temperature-sensitive. Although these do not come under the subject of this lecture, it is as well to refer to them, since they may be misleading. Such extra reflections, which are usually found on photographs of crystals having cleaved, ground or fractured surfaces, and are particularly common in the case of metals, are almost certainly due to the disordering of the lattice in one or more directions, by strain. Such effects the simple diffraction theory is adequate to deal with. In the case of a photographic survey they are unlikely to cause confusion because their sharpness and structure-sensitiveness distinguish them from the extra reflections due to the heat motions. In methods involving ionization measurements they are not so obviously distinguishable (Kirkpatrick, 1941 ; Teague, 1941 ; Lonsdale and Smith, 1941).

Diamond

This crystal will be considered separately because of its peculiarities, which as yet have received no adequate explanation. The elastic constants of diamond are not yet fully known, but they are certainly very large, and it is unlikely that they are very anisotropic. (The compressibility is $\sim 6\cdot10^{12}$ dynes/cm^2 (Adams, 1921, Williamson, 1922), and Young's modulus in the octahedral plane is given as $5\cdot5 \,.\, 10^{12}$ dynes/cm^2 by Pisharoty (1940).) The extra reflections to be expected on account of the elastic vibrations will therefore be single diffuse spots, elliptical in shape, which will fade out quickly as the angle of setting departs from the Bragg angle. Such reflections have been found for all observable planes of all the diamonds examined; they disappear almost completely at liquid-air temperatures (−180° c.) and are increased considerably in intensity when the temperature is raised to 750° c. They correspond to a small ellipsoidal spreading of reflecting power about the reciprocal-lattice points. These extra reflections have been called *primary reflections* because some diamonds also show *secondary extra reflections* of a totally different character (Lonsdale and Smith, 1941 c). It was to these secondary reflections that Raman and Nilakantan drew attention in announcing their new theory. The secondary reflections are relatively sharp and are only slightly, if at all, temperature-sensitive. They are, however, very structure-sensitive, being given in varying intensity by normal (type I) diamonds, but not at all by the rare diamonds classed as being of type II (plate 5 (*c*) and (*d*)) (Robertson, Fox and Martin, 1934; Lonsdale, 1942). Normal diamonds, although monolithic, are generally believed to be in a state of internal strain, whereas in the rarer type II diamonds it is supposed that the strain has been relieved by the breaking up of the diamond into a mosaic of perfect crystallites. The secondary reflections found for {111} planes correspond to a sharply defined extension of reflecting power along all the cube directions. This extension leads to an interesting breakdown of Friedel's law; the reflection from a (111) plane changes when the plane is rotated through 180°, relative to the incident beam, that is, when the angle of incidence changes from θ_B to $-\theta_B$. This is because the extension of reflecting power is not symmetrical about the radius vector of the (111) reciprocal-lattice point. The extension along the cube directions about the {111} points in reciprocal space is *geometrically* accounted for either by the existence of sheets of atoms scattering in phase, parallel to the cube planes, but with a randomness of phase in other directions (W. H. Bragg, 1941), or by the existence of sharply defined lattice waves propagated only along cube directions, the atomic movements being normal or nearly normal to the {111} planes (Pisharoty, 1941; Lonsdale and Smith, 1941; Sirkar and Bishui, 1941; Lonsdale, 1942). Neither of these hypotheses, however, satisfactorily accounts for the fact that the extension of reflecting power about the (220) and (113) planes is not along *all* the cube directions, but only along [100] and [010]; the [001] intensity spike is missing for those two planes, but is present for the (331) plane. The (331) and (004) planes show no intensity spikes along [010] and [100]. The absence of certain intensity spikes could have been accounted for in terms of polarization of certain lattice waves, but the particular combinations that are found do not fit in with each other on any such hypothesis; while the idea of simple two-dimensional sheets of atoms does not permit the absence of any intensity spikes at all.

Bearing these unexplained experimental facts in mind, it is interesting to consider the third theory that has been advanced.

§ 6. RAMAN THEORY

Raman had rejected the Faxén-Waller theory on the erroneous assumption that it inevitably leads to a very weak spherical distribution of intensity around each reciprocal-lattice point (including the origin) which is proportional to the Bragg intensity for each point.

He has suggested instead that the vibrations responsible for the extra reflections must be monochromatic vibrations in the optical branches, of frequency 10^{12} to 10^{14} sec.^{-1}, which are, he suggests, excited by the incident x rays. In order to account for the secondary reflections from (111) diamond planes (the observations made by Raman and Nilakantan do not appear to have included primary reflections, nor reflections from any other planes of diamond) he suggests that the 1332 cm.^{-1} vibration, in which neighbouring carbon atoms vibrate against each other normal to the (111) planes, is excited by the x rays all over the crystal simultaneously, but that there are small changes of phase from one unit cell to the next which are of such a nature that "phase-waves" are propagated along the cube directions. These phase-waves, whose frequency is determined by the frequency of the characteristic vibration, reflect the x rays coherently, with a change of frequency equal to that of the excited vibration. (Since the excited vibration is of a frequency far smaller than that of the incident radiation, this change would be quite impossible to measure. A similar change is given by the Faxén-Waller theory, as was first pointed out by Laue, but it is of little experimental significance, except in so far as it leads to the essential incoherence, and therefore diffuseness, of the extra reflections.)

Raman gives for the law of temperature variation of the vibration energy a factor $\frac{e^{h\nu/kT}+1}{e^{h\nu/kT}-1}$, which is of the same form as that given in the Faxén-Waller theory, although his method of deduction is not clear. There is this important difference in practical application, however: that whereas in the Raman theory only the excited monochromatic frequencies should be used in this formula, according to the Faxén-Waller theory the frequencies or range of frequencies which must be considered are those which give the particular extra spots under observation.

Venkateswaran (1941) has made a few measurements of the variation of intensity of extra reflections with temperature, on carborundum, $NaNO_3$, NaCl, and pentaerythritol, and has suggested that they may be fitted to the formula $I_T \propto \frac{e^{h\nu/kT}+1}{e^{h\nu/kT}-1}$ if ν corresponds to about 800 cm.$^{-1}$ for carborundum, 200 cm.$^{-1}$ for $NaNO_3$, 160 cm.$^{-1}$ for NaCl and <100 cm.$^{-1}$ for pentaerythritol. These frequencies he identifies with known *Reststrahlen* or Raman frequencies. His experiments and deductions may be criticized on the following grounds:

(1) For $NaNO_3$ and for $C(CH_2OH)_4$ he has only two measurements, one at 90° K. and another at 300° K. These, being relative intensities, cannot be used to determine any curve uniquely. Baltzer (1941), for instance, has obtained relative intensity measurements for NaCl crystals at 300° A. and 100° A., which are numerically in close agreement with the measurements of Venkateswaran, but he has fitted them to the formula $I_{100} = I_{300}(1-e^{-2M})$ given by Jauncey, Harvey and Woo (1931, 1932).

(2) He has neglected the effect of temperature in modifying the atom-scattering factor. The spread of electron density due to atomic vibration results in a decrease of the effective atom-scattering factor, which is greater the higher the temperature, the smaller the spacing and the larger the amplitude of vibration. By measurement of this factor for Bragg reflections of NaCl, Waller and James (1928) have been able to show that the root-mean-square amplitude of the Cl atoms is 0·217 A. and of the Na atoms 0·242 A., both at 290° K. For the 200 and 220 reflections (which Venkateswaran presumably measured) this corresponds to a decrease of intensity of 16 and 28 per cent. respectively, at 300° K., as compared with 90° K. This modification of the f-factor to fe^{-M} (where M is a function of temperature, total angle of deflection, wave-length and the interatomic forces) applies also to the extra reflections (Born and Sarginson, 1941), and it is important to notice that it results in a decrease of intensity with increasing temperature. Allowing for this factor, the real increase in the intensity due to the energy factor $\frac{e^{h\nu/kT}+1}{e^{h\nu/kT}-1}$ will be considerably larger than the measured increase, and the deduced value of ν will be considerably smaller. In the case of NaCl, which is the only

one of the four compounds for which a numerical estimate of the necessary correction is available, the appropriate value of ν would correspond to about 100 cm.$^{-1}$ rather than 160 cm.$^{-1}$. This is, in fact, *below the range of optical frequencies for NaCl* (Kellermann, 1940). It corresponds to an acoustical frequency $\nu \doteq 3 \,.\, 10^{12}$ of rather short wave-length.

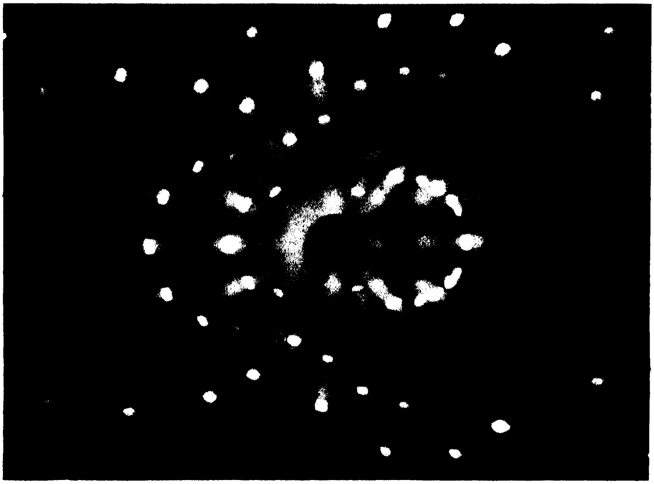
(3) In order to measure the intensity of the extra reflections at all, Venkateswaran has had to choose spots which are separated from neighbouring Laue spots by a region of low background intensity. This means that his measurements are necessarily made on extra reflections of fairly large wave-vector, that is, due to vibrations of fairly short wave-length. For instance, in the case of NaCl, it may be seen *by a geometrical construction* that the vibrations corresponding to the 200 and 220 extra reflections on his photographs are of wave-lengths of only about 22·5 A. and 12 A. respectively. Such waves must have relatively high frequencies in any case. All that Venkateswaran's experiments did in fact show was that the frequencies corresponding to relatively short waves would be of decreasing order of magnitude for SiC, $NaNO_3$, NaCl and $C(CH_2OH)_4$, a result which is fully in accordance with the Faxén-Waller theory, since that is the order of decreasing strength of quasi-elastic forces in the crystals.

Raman has argued that the x-ray diffraction due to the elastic waves would be proportional to N, the number of unit cells in the crystal; but that that due to excited characteristic waves would be proportional to N^2, and, therefore, of the same order as that of the normal Bragg reflections. There seems here to be a misunderstanding of ordinary diffraction theory. The integrated Bragg reflection from a small crystal is certainly proportional to $N_0^2 \delta V$ (*International Tables for Crystal Structure Determination*, **2**, 562 (1935)), but in this case N_0 is the number of unit cells per unit volume and N_0^2 is a constant for any given compound. The volume of the small crystal, δV, is proportional to the number of unit cells in the crystal, and, therefore, *the integrated intensity of the Bragg reflection*—like that of the extra reflections due to elastic vibrations—*is proportional to the number of unit cells in the crystal, N.* All the experiments carried out by Raman and his colleagues to prove that the Bragg and extra reflections depend in the same way upon the crystal (or irradiated) volume simply confirm the Faxén-Waller theory, because in all of them it is integrated intensities that are measured or estimated, not the intensity at a point.

From the point of view of the accepted theory of lattice dynamics (due to Born and his collaborators) there is no justification for Raman's differentiation between the acoustical and optical vibrations; but Raman rejects the accepted theory and, in particular, he challenges the Born theory of the cyclic lattice. This point is dealt with in a forthcoming paper by Professor Born *, and need not be further considered here. It must be emphasized, however, that from the experimental side there is no evidence to show that the extra reflections due to the optical branches of the frequency spectrum will be of a higher order of intensity than those due to the acoustical branches. The relative intensities are simply controlled, as previously explained, by the relative frequencies, being, in fact, much smaller for the optical vibrations on account of their higher frequencies. In order to overcome this difficulty Raman makes assumptions

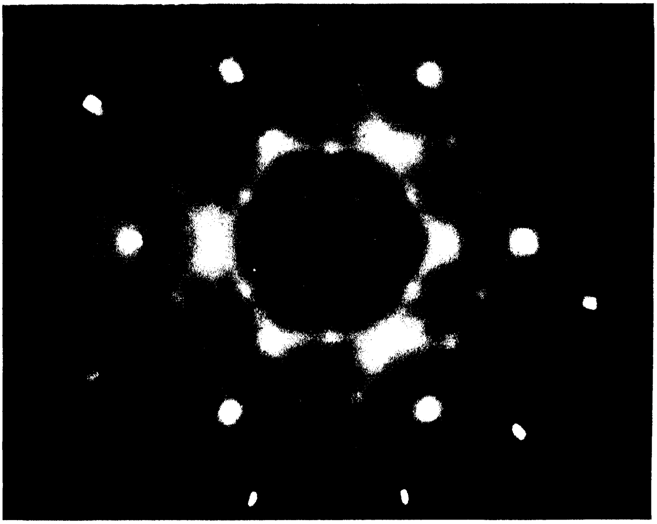
---

* Now published (this part, p. 362).



(*a*)

Sylvine (KCl), showing sharp Laue spots, diffuse extra spots and vertical streaks due to Ag characteristic radiation, and diffuse radial streaks due to white radiation.



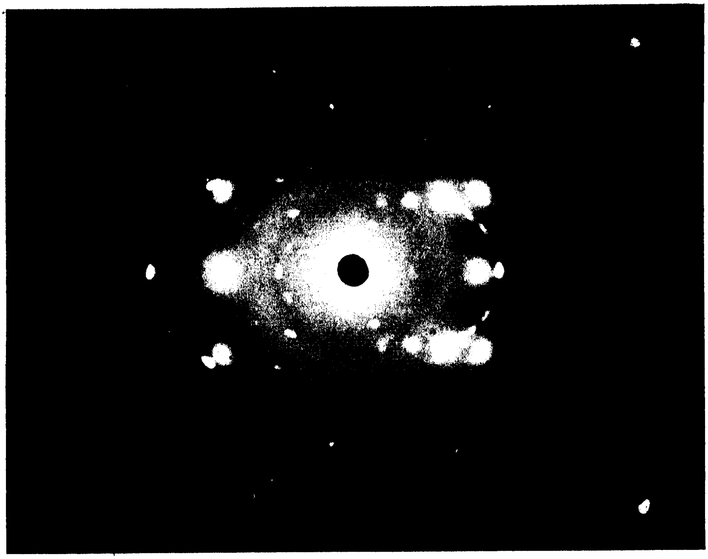
(Reproduced by courtesy of the Council of the Royal Society)

(*b*)

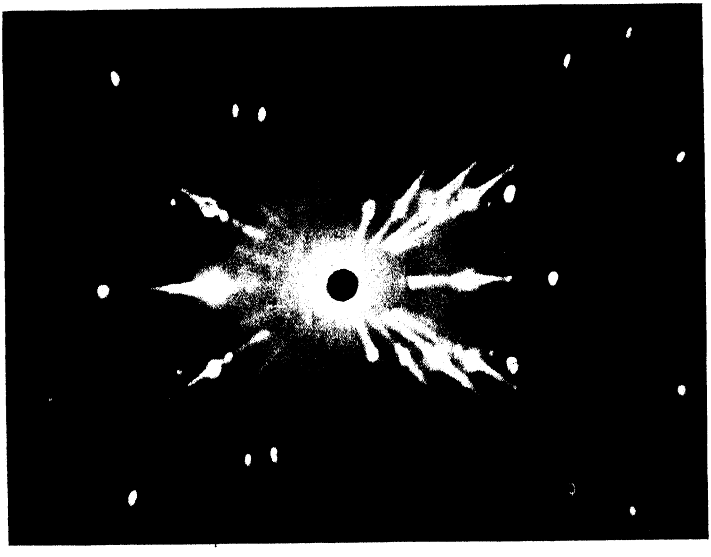
Benzil. Laue photograph taken with monochromatized Cu radiation. No radial streaks, very little background near centre, but strong extra spots and streaks corresponding to intersection of sphere of reflection with certain reciprocal-lattice axes and planes. 2-hr. exposure.

Plate 1.

To face page 342

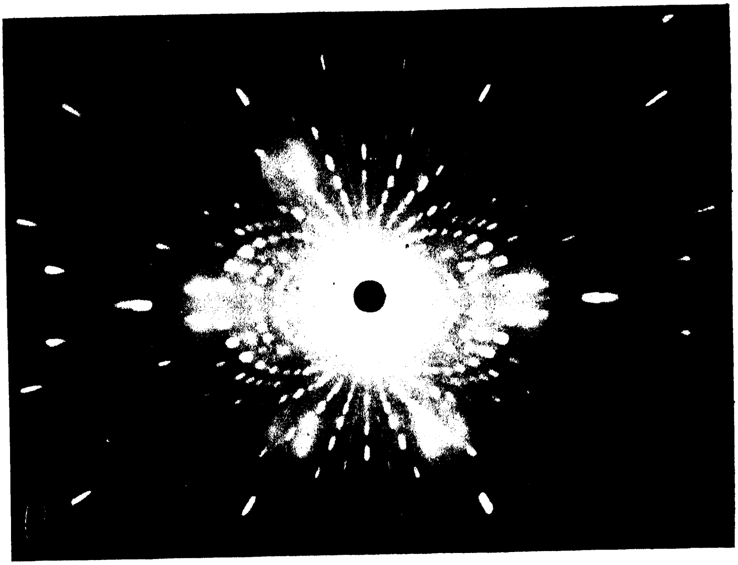


(*a*)
Diphenyldiacetylene; Cu radiation, partly filtered; Laue photograph (crystal stationary), 10-min. exposure. Compare positions of diffuse spots with oscillation spots in (*b*).



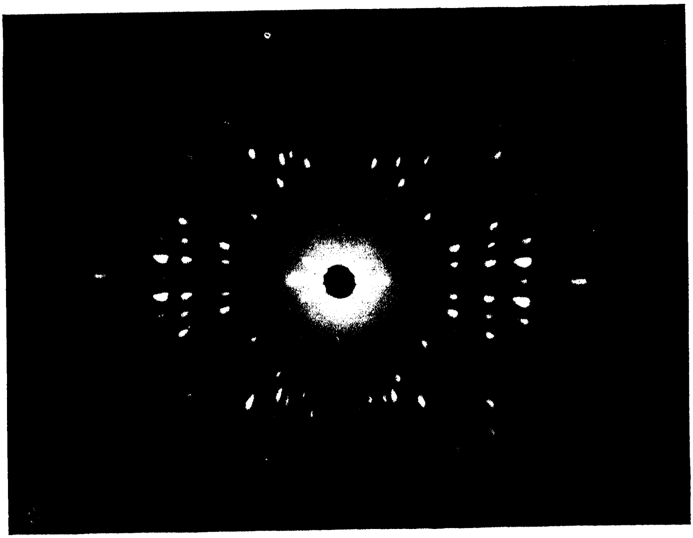
(*b*)
As in (*a*), but crystal oscillating; 10-min. exposure, showing strong "white" radiation streaks and diffuse background to some Bragg spots. (Cu radiation contains some tungsten lines.)

Plate 2.



(*a*)

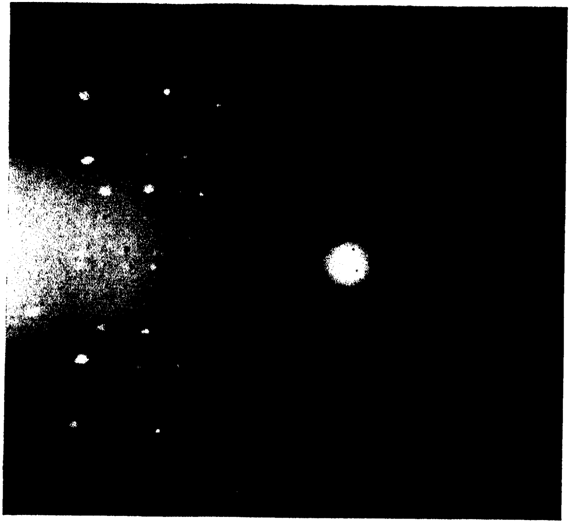
s-**Triphenylbenzene.** This orthorhombic crystal is composed of pseudo-trigonal molecules arranged in approximately parallel orientations. The extra reflections show the pseudo-hexagonal nature of the elastic forces. (Filtered Cu radiation, 20 min.)



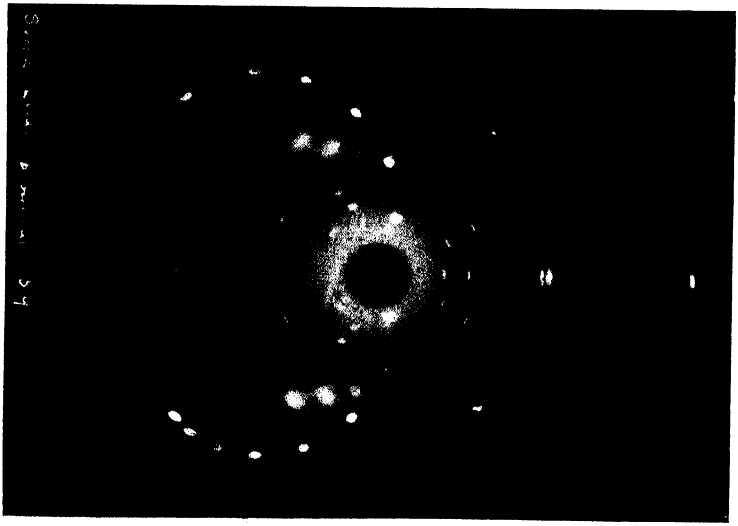
(*b*)

15° oscillation photograph corresponding to above setting of crystal. A careful comparison of these two photographs shows clearly that the intensities of the extra reflections are not necessarily proportional to those of the Bragg reflections. (Filtered Cu radiation, 4 min.)

Plate 3.

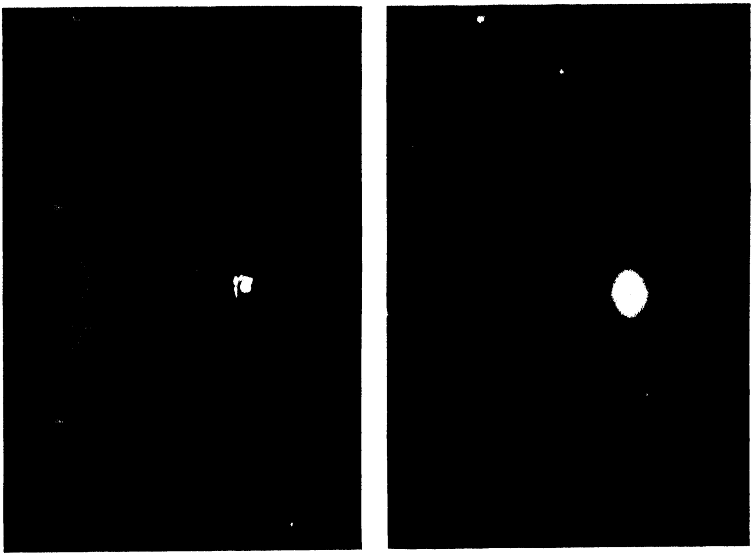


(*a*)
Typical *layer structure*: graphite 006 reflection at room temperature. Cu radiation. One large extra reflection.



(*b*)
Typical *chain structure*: sorbic acid. Streaky extra reflections correspond to planes nearly normal to chain axis; intense elliptical extra reflections correspond to planes nearly parallel to plane of molecules.

Plate 4.

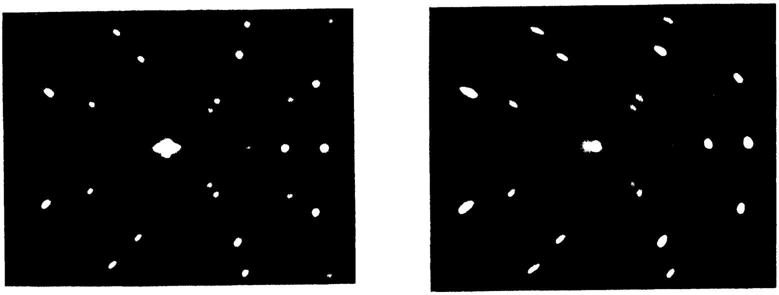


(*a*) (*b*)

(*a*) Graphite 004 reflection on Laue photographs at −180° c.
(*b*) ,, ,, ,, ,, at 25° c.,
showing disappearance of diffuse reflection at low temperatures.

Diamonds

Type I (normal) Type II (rare)

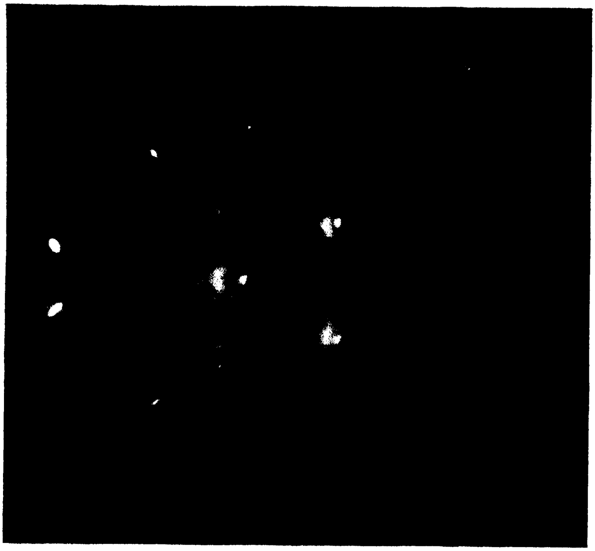


(*c*) (*d*)

(*c*) Diamond type I. The triangles of secondary extra spots associated with the $11\bar{1}$ Laue reflection are given by FeKα (strong) and FeKβ (weak) radiations respectively.

(*d*) Diamond type II. Similar crystal orientation. Only a single primary extra spot, due to FeKα radiation, is to be seen just by the $11\bar{1}$ Laue reflection.

Plate 5.



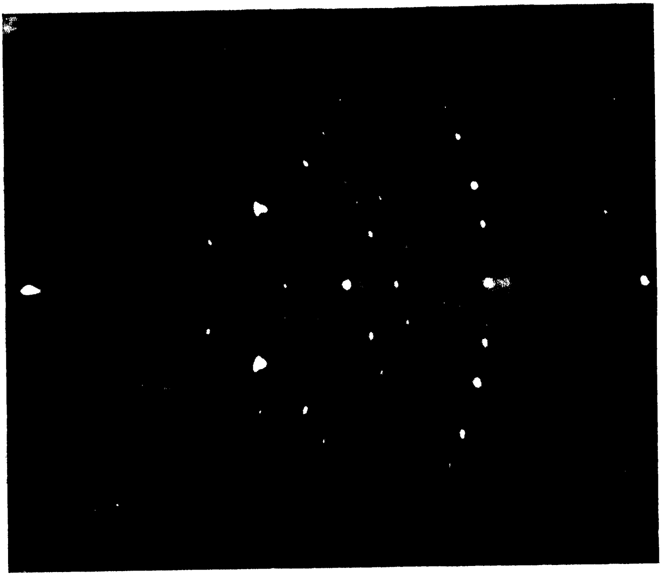
(*a*)

Sodium: Cu radiation, showing 002 extra reflection shaped like a pair of scissors. Note relative weakness of hinge of scissors and presence of strong non-radial streaks.

(*b*)

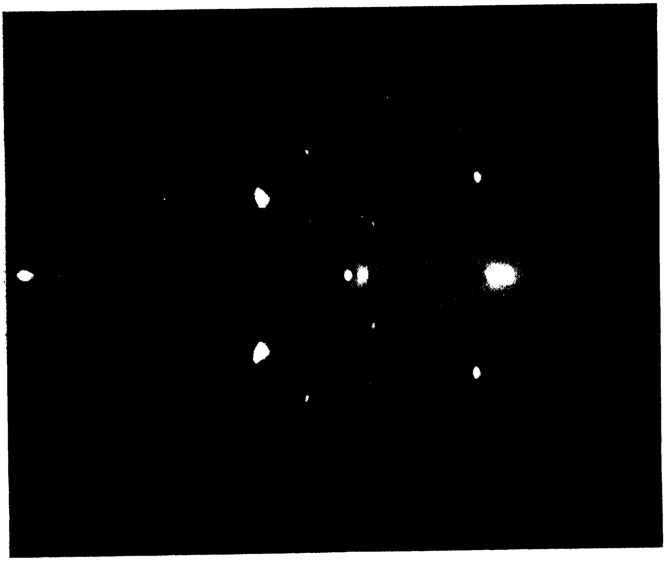
Sodium: Cu radiation, 110 extra reflection, showing strong vertical streak through close group of four spots forming square. (These four spots are clearly resolved in the original film, but detail has been lost in the reproduction.)

Plate 6.



(*a*)

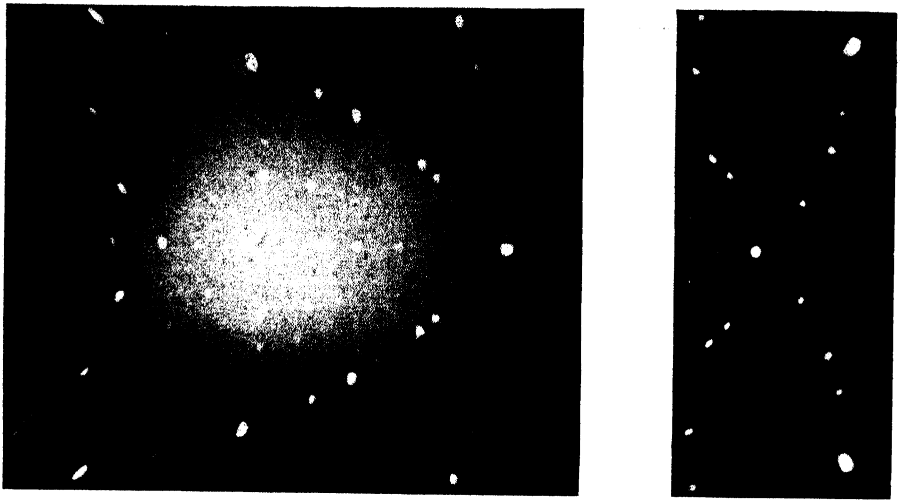
Lead: Cu radiation; Laue photograph at −180° c.; cube axis vertical; film parallel to incident beam; extra reflections much weaker than in (*b*), but many more Laue spots present.



(*b*)

Lead: Laue photograph at room temperature, showing strong extra reflections, which are very persistent as crystal is moved out of position. Exposure, 15 min. at 3·0 cm. in both photographs.

Plate 7.



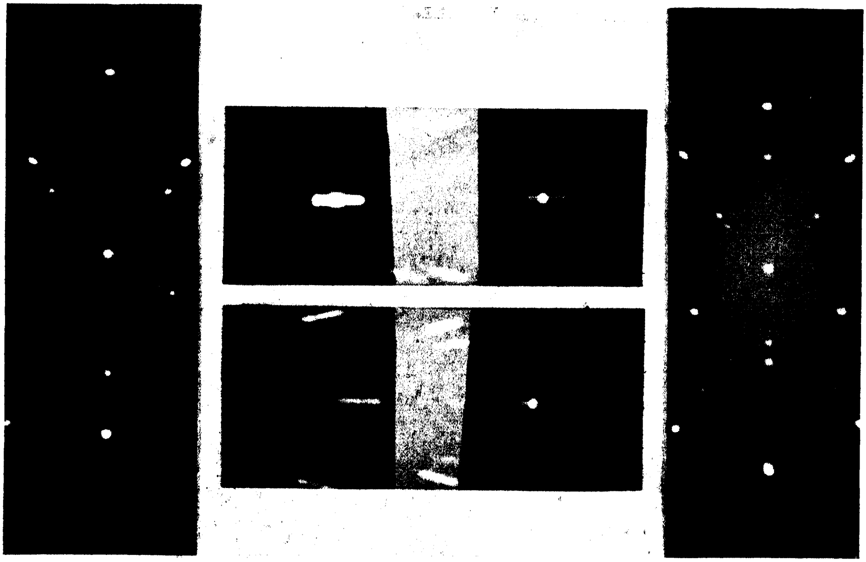
(a)

(b)

(a) Type I diamond, showing 202 reflection with large primary diffuse spot and sharp secondary spot corresponding to intersection of sphere of reflection with [001] horn. No secondary spot corresponding to a [010] horn (the [100] spot would not be observable in this orientation). Similarly, 022 reflection gives only [001] secondary spot, not [100].

(b) Type I diamond, showing 111, 202 and 022 reflections (last two in Bragg positions).

(d)



(c)

(e)

(f)

Type I diamond.

(c) $22\bar{2}$ Laue reflection ($\theta_{\text{Bragg}} - 1^\circ$), showing no secondary detail.

(d) $11\bar{1}$ oscillation, showing secondary streaks in background, compared with $22\bar{2}$ oscillation (same exposure time and development), showing no streaks except that due to white radiation.

(e) Oscillation through 200 position, showing no characteristic reflection, compared with $22\bar{2}$ oscillation under similar conditions.

(f) 222 Laue reflection ($\theta_{\text{Bragg}} - 3^\circ$) showing no secondary detail.

Plate 8.

which, as Born has pointed out, contradict the elementary laws both of classical and of quantum mechanics, and lay a quite unjustifiable emphasis upon the optical range of frequencies.

Forbidden reflections

The distinctive feature of the Raman theory is, of course, the suggestion that the vibration energy is derived from the incident x rays. The possibility of an excitation of characteristic vibrations by x-ray energy was considered by Ott in 1935, and has recently been reconsidered by Born, both of whom have come to the conclusion that such excitation may certainly take place, but that the x-ray interference effects due to such high-frequency vibrations will be negligible compared with those due to the low-frequency elastic vibrations, since the intensity of x-ray diffraction varies inversely as the $\overline{\text{frequency}^2}$. For the acoustical vibrations, $\overline{\nu^2}$ is a function of μ which becomes zero when $\mu = \infty$, a, $a/2$, $a/3 \ldots$ (a being the lattice constant); for the optical vibrations ν^2 is never zero, but has minimum values for $\mu = 2a$, $2a/3$, $2a/5 \ldots$ (Born, 1942). These minimum values may be lower than some of the high-frequency acoustical vibrations, but they occur at the edge of the Brillouin zone and correspond to extra reflections about half-way between neighbouring reciprocal-lattice points. It is only in these positions that the infra-red vibrations may be observable in some cases. Observations in this region are therefore of considerable interest.

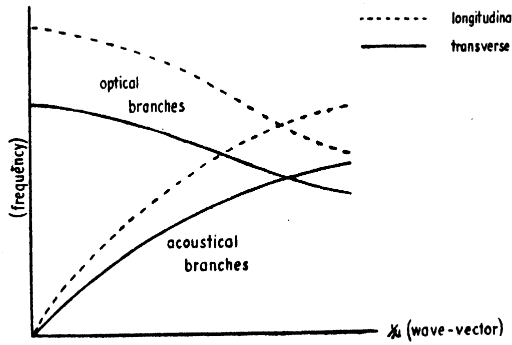


Figure 24.

Ott (1935; footnote to p. 190) suggested that it was possible that a " Raman effect " might be observable in, say, the 222 reflection position for diamond, and might possibly account for the well-known " forbidden " reflection found in that position. This suggestion has been revived by Raman and by one of his collaborators, P. Rama Pisharoty (1941), who has given a list of " permitted " and " forbidden " Bragg and extra reflections in diamond, on the basis of a theory which assumes the excitation of characteristic 1332 cm^{-1} vibrations perpendicular to one set of (111) planes only. This procedure is most arbitrary, for whatever oscillations take place are certainly symmetrical with respect to all the {111} planes, all of which show similar extra reflections. It is also misleading, for vibrations normal to (111) planes would certainly have a different component relative to (220) planes, say, from those normal to $(\bar{1}11)$; but this he has not taken into account. He claims that there should be no extra reflections for (220) planes of intensity at all comparable with those found for (111), and states that Nilakantan, in spite of many trials, has failed to find (220) extra reflections. This is quite at variance with the observations of Lonsdale and Smith (1941 c: cf. plate 8 (*a*) (*b*)), who have found 220 extra reflections of appreciable intensity in many different orientations. Lonsdale and Smith (1942 b) have also tested Pisharoty's other predictions from the Raman theory, namely, that the "forbidden" 222 reflection is an extra and not a Bragg reflection, and that there should be a 200 extra reflection of nearly the same intensity, and have found both to be quite incorrect. There is no 200 reflection of intensity even as much as 1 per cent of the observed 222 reflection, and the 222 reflection has all the characteristics of a normal Bragg reflection and none of those of an extra reflection (plate 8 (*c*)–(*f*)).

One word of caution may be interpolated here. In looking for possible reflections in "forbidden" positions, or at the edge of the Brillouin zone, it is essential that great care should be taken to ascertain whether any supposed effect is really due to the characteristic wave-length from the tube, or whether it may not be due to a sub-multiple of that wave-length. As mentioned before, even the most carefully monochromatized radiation contains unwanted components, which can quite easily amount to 1 per cent or more of the incident intensity. Lonsdale and Smith did indeed find a reflection in the 200 position (<1 per cent of the 222 intensity), but it was certainly due, mainly if not entirely, to a reflection by (400) planes of unwanted $\lambda/2$ radiation, of which the proportion present in the incident beam would only need to be 0·03 per cent in order to account for it completely.

§7. CRUCIAL EXPERIMENTS AND LIMITATIONS OF THEORY

Observations of the extra reflections from single crystals of metals have been most useful in deciding between the theories put forward. The breakdown of the simple diffraction theory in respect of these crystals has been indicated on p. 326. The Raman theory cannot explain the extra reflections from metals at all, except on the artificial assumption of infra-red characteristic vibrations of large amplitude unpredicted by normal lattice theory, and of phase-waves whose direction is unknown except in terms of the elastic forces whose importance the theory denies. The Faxén-Waller theory, as developed by Born, Sarginson and Jahn, explains the detail of the extra reflections from such metals as have been examined very simply and convincingly on the basis of the relative values and anisotropies of the elastic constants. In fact the only criticism that can be made is that the experimental and theoretical agreement found is almost too good. For the thermal theory is still only in the first stage of development, and assumptions are made in it whose validity is not altogether certain over the whole range in which it has apparently given satisfactory agreement with experiment. It is assumed that the amplitudes of vibration are small relative to the crystal spacings and that the effect of displacements of neighbouring atoms can be neglected. Let us briefly examine this point.

The mean K.E. of a vibrating atom (treated as a Planck oscillator) is given by

$$\frac{1}{2}m\overline{\xi^2}(2\pi\nu)^2 = h\nu\left(\frac{1}{2}+\frac{1}{e^\beta-1}\right).$$

Hence
$$\overline{\xi^2} = \frac{h}{2\pi m(2\pi\nu)} \cdot \frac{e^\beta + 1}{e^\beta - 1},$$

where β is $h\nu/kT$, m the actual mass of the atom, $\overline{\xi^2}$ the mean-square amplitude, $2\pi\nu$ the angular frequency, T absolute temperature and h and k the Planck and Boltzmann constants.

Assuming the vibration to be resolvable along three principal directions, the $\overline{\xi_x^2}$ normal to a single plane will be

$$\overline{\xi_x^2} = \frac{h}{6\pi m(2\pi\nu)} \cdot \frac{e^\beta + 1}{e^\beta - 1},$$

and this may be taken as the mean-square amplitude corresponding to a single transverse or longitudinal vibration. The limits of $2\pi\nu$ for the normal modes of vibration, and the values of $2\pi\nu$ corresponding to their maxima in the frequency spectrum of NaCl, are given by Kellermann (1940, 1941) as follows:—

(1) transverse acoustical: $0 \to 3{\cdot}0 \,.\, 10^{13}$ sec.$^{-1}$ (max. $N(\nu)$ at $1{\cdot}9 \,.\, 10^{13}$ sec.$^{-1}$);

(2) longitudinal acoustical : $0 \to 3{\cdot}5 \,.\, 10^{13}$ sec.$^{-1}$ (max. $N(\nu)$ at $3{\cdot}0 \,.\, 10^{13}$ sec.$^{-1}$);

(3) transverse optical: $2{\cdot}3 \,.\, 10^{13} \to 3{\cdot}6 \,.\, 10^{13}$ sec.$^{-1}$ (max. $N(\nu)$ at $2{\cdot}85 \,.\, 10^{13}$ sec.$^{-1}$);

(4) longitudinal optical : $3{\cdot}4 \,.\, 10^{13} \to 6{\cdot}0 \,.\, 10^{13}$ sec.$^{-1}$ (max. $N(\nu)$ at $4{\cdot}1 \,.\, 10^{13}$ sec.$^{-1}$).

Taking for m a mean atomic mass, the corresponding values of $\sqrt{\overline{\xi_x^2}}$ for a Na or Cl ion (considered as equivalent) will be, at 290° K.:—

(1) transverse acoustical: $\infty \to 0{\cdot}080$ A. (for max. $N(\nu)$, $0{\cdot}125$ A.);

(2) longitudinal acoustical: $\infty \to 0{\cdot}069$ A. (for max. $N(\nu)$, $0{\cdot}080$ A.);

(3) transverse optical: $0{\cdot}103 \to 0{\cdot}067$ A. (for max. $N(\nu)$, $0{\cdot}084$ A.);

(4) longitudinal optical: $0{\cdot}071 \to 0{\cdot}043$ A. (for max. $N(\nu)$, $0{\cdot}060$ A.).

The larger values of the acoustical amplitudes and the smaller values of the optical amplitudes will correspond to points near to the Bragg points in reciprocal space, the smaller values of the acoustical amplitudes and the larger values of the optical amplitudes to points half-way between these Bragg points. It will be seen that even the *minimum* values of $\sqrt{\overline{\xi_x^2}}$ are not negligible compared with the nearest distance between neighbouring atoms (2·814 A. from Na to Cl).

The very important X-ray measurements of James, Firth and Waller (1928) confirm the above calculations. They found 0·140 A. and 0·125 A. for the *average* values of $\sqrt{\overline{\xi_x^2}}$ for Na and Cl respectively in the NaCl crystal at room temperatures, giving a mean value which is very near to the "transverse acoustical" maximum $N(\nu)$ peak amplitude. If the average $\sqrt{\overline{\xi_x^2}}$ is as large as 0·13 A., it is fairly clear that much larger amplitudes must exist.

From Jahn's (1942) formulae it is also clear that for elastically anisotropic crystals the acoustical frequencies are lower, and, therefore, the values of $\sqrt{\overline{\xi_x^2}}$ are larger for waves travelling in some directions than in others. Calculations based on known elastic constants (assuming wave velocity to be independent of wave-length) indicate that for transverse vibrations travelling along certain main crystallographic planes or axes, the amplitudes may be very much larger than the measured or calculated average, even for waves of short length. This is confirmed by the presence of non-radial extra streaks on photographs of KCl, benzil, Na, etc. (plates 1 and 6), which signify the existence of vibrations of very large amplitude travelling in the directions given by the streaks, the vibration wave-lengths having all values from infinity down to a length comparable with the unit-cell size. If these large vibration amplitudes were of the optical modes, neighbouring atoms or molecules moving in *opposite* directions, enormous crystal forces would be brought into action to resist the changes in inter-atomic distance thus caused. Short-wave acoustical vibrations of large amplitude could occur with far less strain on the crystal lattice, the *relative* movements of neighbouring atoms being small. The Faxén-Waller theory as it stands is not intended to apply in the short-wave regions, but since many of the most interesting experimental observations are made in that region, and since it does seem to be capable of application there, an extension of the theory is obviously desirable.

In view of the existence of these large amplitudes of vibration, even at room temperatures, and still more at high temperatures, it is not surprising that atoms occasionally break loose, as it were, and either interchange places with other atoms in the lattice or wander from one vacated place to another until they can reach a fault or a channel through which they can move still more easily. From this point of view, the problem of self-diffusion in metals or of electrical conductivity (forced diffusion) in ionic crystals is intimately related to the general dynamics of the lattice.

§ 8. CONCLUSION

It will be realized from the necessarily condensed account given above that this is a subject which is capable of expansion in many directions.

The observations already made form a direct experimental confirmation of the theory of lattice dynamics, and further study of the extra reflections for substances of simple structure should lead to better knowledge of the quasi-elastic forces between the atoms and molecules in the solid state, and of the changes in those forces with change of temperature.

A secondary method of crystal-structure determination is now available, to supplement the direct methods of Patterson-Harker and Fourier analysis. Conversely, it may be possible, using crystals of known structure, to study the atomic vibrations, particularly those of layer and chain structures, which present features of especial interest.

At present the need on the experimental side is for more accurate and comprehensive intensity measurements, and on the theoretical side for a more precise and detailed investigation of the X-ray diffraction at positions in reciprocal space not near to the lattice points, for large angles of deviation and at high temperatures, with proper allowance for the far from negligible amplitudes of vibration of each atom and of its neighbours.

§ 9. ACKNOWLEDGEMENTS

In preparation for this lecture I have received much inspiration from discussions with Sir William Bragg and Professor Max Born. I have also been privileged to consult Dr. Jahn's manuscripts before publication, and am grateful to him for this opportunity. My colleague, Mr. H. Smith, has spared neither time nor trouble in helping me to prepare suitable photographs for purposes of illustration.

REFERENCES

Adams, L. H., 1921. *J. Acad. Sci., Wash.*, **11**, 45.

Baltzer, O. J., 1941. *Phys. Rev.* **60**, 460.

Bernal, J. D., 1926. *Proc. Roy. Soc.* A, **113**, 117.

Born, M., 1941. *Nature, Lond.*, **147**, 674 ; 1942. *Ibid.* **149** 403 ; *Proc. Roy. Soc.* (in press).

Born, M. and Göppert-Mayer, M., 1933. *Handb. d. Physik*, **24**/2, 638.

Born, M. and v. Kármán, Th., 1912. *Phys. Z.* **13**, 297 ; 1913. *Ibid.* **14**, 15, 65.

Born, M. and Sarginson, K., 1941. *Proc. Roy. Soc.* A, **179**, 69.

Bragg, Sir W. H., 1940. *Nature, Lond.*, **146**, 509 ; 1941. *Ibid.* **148**, 780 ; *Proc. Roy. Soc.* A, **179**, 51, 94.

Bragg, Sir W. L., 1941. *Proc. Roy. Soc.* A, **179**, 61.

Brill, R., Grimm, H. G., Hermann, C. and Peters, Cl., 1939. *Ann. Phys., Lpz.*, **34**, 393.

Clark, G. L. and Duane, W., 1922. *Proc. Nat. Acad. Sci., Wash.*, **8**, 90 ; 1923. *Ibid.* **9**, 126, 131 ; *J. Opt. Soc. Amer. and Rev. Sci. Instrum.* **7**, 455 ; *Science*, **58**, 400 ; 1924. *Proc. Nat. Acad. Sci.* **10**, 48.

Coven, A. W., 1932. *Phys. Rev.* **41**, 422.

Debye, P., 1914. *Ann. Phys., Lpz.*, **43**, 49.

Duane, W. and Clark, G. L., 1923. *Phys. Rev.* **21**, 379.

Ewald, P. P., 1921. *Z. Kristallogr.* **56**, 129 ; 1940. *Proc. Phys. Soc.* **52**, 167.

Faxén, H., 1923. *Z. Phys.* **17**, 266.

Friedrich, W., 1913. *Phys. Z.* **14**, 1082.

Gregg, R. Q. and Gingrich, N. S., 1941. *Phys. Rev.* **59**, 619.

Hughes, E. W., 1941. *J. Amer. Chem. Soc.* **63**, 1737.

Jahn, H. A., 1941. *Nature, Lond.*, **147**, 511 ; 1942. *Proc. Roy. Soc.* A, **179**, 320.

Jahn, H. A. and Lonsdale, K., 1942. *Phys. Rev.* **62**, 355.

James, R. W. and Firth, E. M., 1928. *Proc. Roy. Soc.* A, **117**, 64.

James, R. W., Waller, I. and Hartree, D. R., 1928. *Proc. Roy. Soc.* A, **118**, 334.

Jauncey, G. E. M., 1922. *Phys. Rev.* **20**, 405, 421.

Jauncey, G. E. M. and Baltzer, O. J., 1940. *Phys. Rev.* **58**, 1116 ; 1941. *Ibid.* **59**, 699.

Jauncey, G. E. M., Baltzer, O. J. and Miller, D. C., 1941. *Phys. Rev.* **59**, 908.

Jauncey, G. E. M. and Harvey, G. G., 1931. *Phys. Rev.* **37**, 1203.

Jauncey, G. E. M. and May, H. L., 1924. *Phys. Rev.* **23**, 128.

Jauncey, G. E. M. and Pennell, F., 1933. *Phys. Rev.* **43**, 505.

Kellermann, E. W., 1940. *Philos. Trans.* A, **238**, 513 ; 1941. *Proc. Roy. Soc.* A, **178**, 17.

Kirkpatrick, P., 1941. *Phys. Rev.* **59**, 699.

Kracek, F. C., Posnjak, E. and Hendricks, S. B., 1931. *J. Amer. Chem. Soc.* **53**, 3339.

Laue, M. v., 1926. *Ann. Phys., Lpz.*, **81**, 277 ; 1927. *Z. Kristallogr.* **65**, 493 ; 1936. *Ann. Phys., Lpz.* **26**, 55.

Laue, M. v. and Riewe, K. H., 1936. *Z. Kristallogr.* **95**, 408.

Laval, J., 1938. *C.R. Acad. Sci., Paris*, **207**, 169 ; 1939. *Ibid.* **208**, 1512 ; *Bull. Soc. Franç. Min.* **62**, 137.

Lonsdale, K., 1942. *Proc. Roy. Soc.* A, **179**, 315.

Lonsdale, K., Robertson, J. M. and Woodward, I., 1941. *Proc. Roy. Soc.* A, **178**, 43.

Lonsdale, K. and Smith, H., 1941 a. *Proc. Phys. Soc.* **53**, 529 ; 1941 b. *Proc. Roy. Soc.* A, **179**, 8 ; 1941 c. *Nature, Lond.*, **148**, 112, 257 ; 1941 d. *Ibid.* **148**, 628 ; 1942 a. *Ibid.* **149**, 21 ; 1942 b. *Ibid.* **149**, 404.

Mauguin, C. and Laval, J., 1939. *C.R. Acad. Sci., Paris*, **208**, 1446.

Müller, A., 1941. *Proc. Roy. Soc.* A, **178**, 227.

OTT, H., 1935. *Ann. Phys., Lpz.*, **23**, 169.
PISHAROTY P. RAMA, 1940. *Proc. Ind. Acad. Sci., Bangalore*, **12 A**, 208 ; 1941. *Ibid.* **14 A**, 56, 377, 434.
PRESTON, G. D., 1939. *Proc. Roy. Soc.* A, **172**, 116 ; 1941. *Ibid.* A, **179**, 1.
RAMAN, Sir C. V., 1941. *Proc. Ind. Acad. Sci., Bangalore*, **13 A**, 1 ; **14 A**, 317, 332 ; 1942. *Proc. Roy. Soc.* A, **179**, 289, 302.
RAMAN, Sir C. V. and NILAKANTAN, P., 1940 a. *Nature, Lond.*, **145**, 667 ; *Curr. Sci.* **9**, 165 ; 1940 b. *Nature, Lond.*, **146**, 523, 686 ; **147**, 118 ; 1940 c. *Proc. Ind. Acad. Sci., Bangalore*, **11 A**, 379, 389, 398 ; **12 A**, 141 ; 1941. *Ibid.* **14 A**, 356.
ROBERTSON, Sir R., FOX, J. J. and MARTIN, A. E., 1934. *Philos. Trans.* A, **232**, 463.
ROBINSON, B. W., 1933. *Proc. Roy. Soc.* A, **142**, 435.
SARGINSON, K., 1942. *Proc. Roy. Soc.* (in press).
SIEGEL, S., 1941. *Phys. Rev.* **59**, 371.
SIEGEL, S. and ZACHARIASEN, W. H., 1940. *Phys. Rev.* **57**, 795.
SIRKAR, S. C. and BISHUI, B. M., 1941. *Science and Culture*, **7**, 314.
TEAGUE, D. S., 1941. *Bull. Amer. Phys. Soc.* **16** (6), 7.
VENKATESWARAN, C. S., 1941. *Proc. Ind. Acad. Sci., Bangalore*, **14 A**, 387.
WADLUND, A. P. R., 1938. *Phys. Rev.* **53**, 843.
WALLER, I., 1923. *Z. Phys.* **17**, 398 ; 1925. *Dissertation*, Uppsala ; 1928. *Z. Phys.* **51**, 213.
WALLER, I. and JAMES, R. W., 1927. *Proc. Roy. Soc.* A, **117**, 214.
WILLIAMSON, E. D., 1922. *J. Franklin Inst.* **193**, 491.
WOO, Y. H., 1931. *Phys. Rev.* **38**, 6 ; 1932. *Ibid.* **41**, 21.
WYCKOFF, R. W. G., 1923. *Amer. J. Sci.* **5**, 277.
ZACHARIASEN, W. H., 1938. *Phys. Rev.* **53**, 844 ; 1939. *Bull. Amer. Phys. Soc.* (Dec. 1) ; 1940. *Phys. Rev.* **57**, 597 ; 1941 a. *Ibid.* **59**, 207 ; 1941 b. *Ibid.* **59**, 766, 860, 909.

DISCUSSION

Sir WILLIAM BRAGG. Mrs. Lonsdale will describe the work which she has been carrying on at the Royal Institution for some time past. It deals with a phenomenon which may have been observed in the early days of x-ray crystal analysis, but has not been thoroughly investigated until recently ; Mrs. Lonsdale's study of it is by far the most complete up to date. These newly examined patterns present a novel method of examining the internal structure of solids. There is a wide field for investigation, and it is likely that important discoveries will be made in it. I should like to add that I have watched Mrs. Lonsdale's work with great interest, and have admired the skill and the untiring perseverance which she has spent upon it.

Dr. W. A. WOOSTER. I should like to ask whether effects resembling those described in the lecture, and attributed to elastic waves associated with the thermal motion of the atoms of the crystal, cannot be obtained by sending supersonic waves through the crystal. Some years ago x rays were reflected from vibrating crystals of quartz and a modification of the reflection was found. Has an explanation of these results been put forward on the basis of the present theory ?

A second point concerns a remark about the future development of this study. The lecturer stated that it might be possible to determine the elastic constants and their temperature coefficients by observations on diffuse spots. Would such observations depend upon a detailed investigation of the variation of density within a number of such spots ? If so, how can variations in the background round a spot due to the imperfect alignment of the constituent mosaic blocks be eliminated ?

Mr. G. D. PRESTON. In listening to Mrs. Lonsdale's lecture I have been impressed by the necessity of obtaining powerful sources of x rays, such as the 5 and 50 kw. tubes at the Royal Institution, for the study of these background reflexions by means of monochromatized beams. The attempt to explain the diffuse patterns by means of diffraction in a small group (the " simple diffraction hypothesis ") was made to avoid the mathematical complexities of the Faxén-Waller theory : it is not surprising that it fails to account for all the details of the patterns, but it is interesting to find that it can predict some of the finer points of the Na patterns, as Sir William Bragg has recently shown (*Nature, Lond.*, **148**, 780 (1941)). Can the question, raised by Sir Lawrence, as to the reason for the accumulation of intensity near the normal reflexions, be answered in this way? The shortest compressional waves of length $2a$ require neighbouring atoms to move towards or away from one another, and it is unlikely that such a motion can have a large amplitude. A longer wave represented by a point close to a normal reflexion could have a proportionately greater amplitude without involving a greater relative displacement of neighbouring atoms. The success of the simple diffraction hypothesis may, I think, be partly due to the fact that it takes into account the relative immobility of neighbouring atoms by postulating a more or less rigid group of atoms. But the idea is evidently much too simple to afford a satisfactory explanation of the diffuse spots.

Arising out of Dr. Wooster's question, I think any imperfection in the crystal will produce background reflexions. Even if the imperfection is not a spatially periodic one, like the elastic temperature waves, it can be resolved into a Fourier series, each component of which will produce an effect between the normal spectra. Great care would be necessary in selecting perfect crystals for measurement of elasticity in the way suggested by Dr. Wooster.

Dr. G. T. Bennett. The calculation may be simplified as follows: From *Proc. Roy. Soc.* A, **179**, 334, the equation for r^2 is the expanded equivalent of

$$\begin{vmatrix} A_{11} & A_{12} & A_{13} & L \\ A_{21} & A_{22} & A_{23} & M \\ A_{31} & A_{32} & A_{33} & N \\ L & M & N & r^2 \end{vmatrix} = 0 \qquad \ldots\ldots(1)$$

or of

$$\begin{vmatrix} \beta+\delta l^2 & \gamma lm & \gamma nl & L \\ \gamma lm & \beta+\delta m^2 & \gamma mn & M \\ \gamma nl & \gamma mn & \beta+\delta n^2 & N \\ L & M & N & r^2 \end{vmatrix} = 0, \qquad \ldots\ldots(2)$$

where $\alpha=c_{11}$, $\beta=c_{44}$, $\gamma=c_{12}+c_{44}$, $\delta=\alpha-\beta=c_{11}-c_{44}$,

or of

$$\begin{vmatrix} \beta/l^2+\epsilon+\gamma & \gamma & \gamma & L/l \\ \gamma & \beta/m^2+\epsilon+\gamma & \gamma & M/m \\ \gamma & \gamma & \beta/n^2+\epsilon+\gamma & N/n \\ L/l & M/m & N/n & r^2 \end{vmatrix} = 0, \qquad \ldots\ldots(3)$$

where $\epsilon=\delta-\gamma$.

If

$$\left.\begin{aligned} \lambda &= \beta/l^2+\epsilon, \\ \mu &= \beta/m^2+\epsilon, \\ \nu &= \beta/n^2+\epsilon, \end{aligned}\right\} \qquad \ldots\ldots(4)$$

the determinant (3) is the eliminant of

$$\left.\begin{aligned} (\lambda+\gamma)x+\gamma y+\gamma z &= L/l, \\ \gamma x+(\mu+\gamma)y+\gamma z &= M/m, \\ \gamma x+\gamma y+(\nu+\gamma)z &= N/n, \end{aligned}\right\} \qquad \ldots\ldots(5)$$

$$r^2=(L/l)x+(M/m)\,y+(N/n)\,z. \qquad \ldots\ldots(6)$$

If $s=x+y+z$, equations (5) become

$$\left.\begin{aligned} \lambda x+\gamma s&=L/l,\\ \mu y+\gamma s&=M/m,\\ \nu z+\gamma s&=N/n, \end{aligned}\right\} \qquad \ldots\ldots(7)$$

whence

$$\left.\begin{aligned} x&=\frac{L/l-\gamma s}{\lambda},\\ y&=\frac{M/m-\gamma s}{\mu},\\ z&=\frac{N/n-\gamma s}{\nu}, \end{aligned}\right\} \qquad \ldots\ldots(8)$$

and by addition

$$s=\Sigma\frac{L/l}{\lambda}-\gamma s\,\Sigma\frac{1}{\lambda},\text{ giving } s, \qquad \ldots\ldots(9)$$

and hence, also, x, y, z from (8) and r^2 from (6).

Dr. H. WILMAN. The phenomena of the diffraction of fast electrons ($\sim$60 kv.) closely parallel those of x-ray diffraction except that there is often greater prominence of secondary and inelastic scattering effects, though it has been realized recently that secondary scattering is also noticeable throughout the whole field of x rays. This parallelism leads to the expectation that diffractions of the type which Mrs. Lonsdale has discussed and illustrated so elegantly should also occur in electron diffraction.

Such diffraction maxima, close to normal spots and comparable in size with them, have so far not been observed clearly in electron diffraction, although some of the patterns from anthracene (Charlesby, A., Finch, G. I. and Wilman, H., 1939, *Proc. Phys. Soc.* **33,** 479) show abnormal maxima which may be related to this kind of extra x-ray diffraction.* On the other hand, a much more diffuse set of prominent diffraction maxima accompanies the normal single-crystal spot pattern in many cases, notably those of molecular crystals. In 1935 I obtained some photographs (see Finch, G. I., Quarrell, A. G. and Wilman, H., 1935, *Trans. Faraday Soc.* **31**, 1050, fig. 32) from thin anthracene crystals, showing spots strengthened in regularly-spaced regions whose spacing and arrangement appeared to correspond to the linked benzene-ring structure of the molecules. Further photographs (Finch, G. I. and Wilman, H., 1937, *Ergebn. exakt. Naturw.* **16,** 353) showed, beside the spot pattern, strong diffuse maxima in regular rows or in parallel bands ; and diffuse bands in transmission patterns were also yielded by long-chain hydrocarbons and their derivatives. A detailed study of anthracene was made in 1939 (Charlesby, A., Finch, G. I. and Wilman, H., *loc. cit.*), and this showed that the diffuse maxima corresponded very closely in position and shape, and qualitatively in relative intensities, with those to be expected from molecules occupying the same orientations as those in the crystal, but diffracting the electron beam independently. From a comparison with the theoretical scattering distribution, it was concluded that the diffuse maxima were best explained as due to thermal vibrations of the practically rigid molecules about their mean positions in the crystal.

While this work was being prepared for publication we learned that Mr. Preston was writing an account (Preston, G. D., 1939, *Proc. Roy. Soc.* A, **172**, 116) of his x-ray work which first showed clearly the temperature dependence of the intensity of the x-ray extra diffractions from non-molecular crystals. The two sets of results appeared to be related, in that the small coherently vibrating atom groups postulated by Preston were analogous to rigid molecules, and Sir William Bragg's more general demonstration (Bragg, W. H., 1941, *Proc. Roy. Soc.* A, **179**, 51, 94), that the positions of the x-ray extra maxima correspond to diffraction from small crystal units consisting of an atom and its nearest neighbours, brings out the analogy very clearly.

Hitherto no maxima of the very diffuse electron-diffraction kind have been obtained in x-ray patterns from molecular crystals in addition to the other extra spots. It appears that the very diffuse maxima, even if present, would mostly fall outside the field of diffraction recorded in x-ray Laue patterns, and their presence or absence is thus not yet demonstrated.

---

* A closer investigation of these abnormal diffractions, and comparison with the x-ray results now available, has shown that they are in fact identical in nature (see Charlesby, A. and Wilman, H., 1942, *Nature, Lond.*, **149,** 411).

That the diffuse electron maxima are not due to a surface effect seems to be shown by the fact that the molecules giving rise to them are oriented just as in the crystal interior, and further, reflection patterns (figure 1) from the smooth flake-like crystals in (one-degree oriented) films grown from solution are essentially similar to the transmission patterns. The mode of dependence of the intensity of these maxima on crystal thickness has not yet been clearly demonstrated. Although the theory (Charlesby, A., Finch, G. I. and Wilman, H., *loc. cit.* 1939) indicated that the ratio of the maximum intensity in a diffuse maximum relative to the intensity maximum in a Bragg spot should decrease with increasing crystal thickness, it seems that the effect of the usually appreciable beam divergence in the case of x-ray diffraction will tend to reduce or annul this decrease, since the intensity maximum in a Bragg spot will then be proportional to the number of atoms in the thickness of the crystal sheet, instead of proportional to its square. Under these conditions the diffuse type of maxima may well also occur in the x-ray patterns for which thicker crystals are used.

Diffuse lines along principal spot rows in electron-diffraction transmission patterns are yielded by some crystals, for example phthalimide (Finch, G. I. and Wilman, H., 1937, *loc. cit.*), anthracene, stearic acid, dodecylalcohol, graphite, molybdenite and cadmium iodide (figure 2). These are probably similar in nature to the lines described by Mrs. Lonsdale, and presumably originate from the strong anisotropy of the crystal relative to elastic waves passing through it. Similar diffuse lines have been observed in some silver patterns (Goche, O. and Wilman, H., 1939, *Proc. Phys. Soc.* **51,** 625), but these may be due to a remnant of the initial twinned structure of the silver layer, not completely removed by heat treatment.

To conclude, it seems to me that any theory of diffraction by crystals undergoing thermal vibrations should be capable of explaining the electron-diffraction phenomena as well as those of x rays ; and I should be very interested to know whether the Faxén-Waller theory can account for the diffuse electron maxima, and whether a more conclusive search for similar x-ray maxima could be made.

Dr. W. H. TAYLOR. I wish to ask Mrs. Lonsdale for information about the use of the new technique in the examination of crystal structures. She has used experimental data obtained from crystals of known structure in order to elucidate the nature and origin of the " extra reflections ". To what extent has this process been inverted in seeking information as to the structure of a crystal not previously examined by standard methods ?

Mr. F. A. BANNISTER. It follows from the work of Laval, Mrs. Lonsdale and Jahn that substances with the same crystal structure, but possessing quite different elastic constants, show marked differences in the distribution of diffuse spots. What does this imply for the two types of diamond photographs recently published by the lecturer ?

The variation in the size of the diffuse spots with temperature might yield a decision between the Faxén-Waller and Bragg theories.

AUTHOR'S reply. Effects resembling those described in the lecture ought, of course, to result from any forced crystal vibrations, but in the case of ultrasonic waves of large amplitude, the wave-lengths will be so large that the ends of the corresponding wave-vectors will lie in the very close neighbourhood of the reciprocal-lattice points, and any extra reflection that occurs will probably result only in a redistribution of intensity within the Bragg spots. Increases in the absolute intensity of the Laue reflections have been reported (Fox and Carr, 1931 ; Nishikawa *et al.*, 1934 ; Klauer, 1935) for quartz crystals vibrating piezoelectrically, but this has been explained as due to a decrease in primary or secondary extinction (Jauncey and Bruce, 1938 ; Fox and Stebbins, 1939). The inference is that the existence of forced vibrations amounts to a change in the " mosaic condition " of the crystal. A change in intensity of the transmitted (central) beam during oscillation of quartz crystals has also been reported (Fox and Fraser, 1935) and denied (Jauncey *et al.*, 1935, 1936 ; Fox and Stebbins, 1939). The continued sharpness of the transmitted beam is to be expected from the Faxén-Waller theory (Weigle and Smith, 1942).

The second point raised by Dr. Wooster is one of considerable interest. I am not convinced at present that *in general* the existence of a mosaic condition, apart from that imposed by the heat vibrations themselves, is proved. The volume of the Faxén-Waller " temporary super-lattices " corresponding to the maxima of the frequency-distribution curve of (say) NaCl is about (10 A.)^3, but there exist also super-lattices of all sizes up to that of the crystal itself. Although these are temporary, they will have the same effect in reducing extinction as would a static breaking up of the crystal. If we suppose, however (as is certainly sometimes the case), that a static mosaic exists, I think that a study of the crystal at different temperatures would distinguish the one effect from the other. Examination of a single photograph would not be sufficient. A detailed investigation of the density variation within all observable extra spots and streaks will give the distribution of extra scattering power in reciprocal space within the limiting sphere, and in particular the distribution about different reciprocal-lattice points of both lower and higher orders. If the mosaic blocks are sufficiently small (say $<$(50 A.)^3) and disoriented (say by several minutes of arc or more) to produce background effects at all, then such extra reflections will be of the kind predicted by the simple diffraction theory, and will be temperature-insensitive. They will also have larger areas at large angles of total deflection. As a rule, however, the constituent blocks of a static " mosaic crystal " are not small enough, nor are they disoriented enough, for background effects to be expected. A most useful diagram showing the influence of crystal imperfection on the integrated intensity of x-ray reflection has been given by Ewald and Renninger (1935). The conditions implied by the question would only exist in the limiting case represented by the lower right-hand corner of the diagram.

I cannot agree with Mr. Preston that care would have to be taken to choose perfect crystals for investigation. Such crystals are not available, and the Faxén-Waller theory, as it stands, does not apply to perfect, but to ideally imperfect, crystals. As far as I know, the effect of thermal vibrations on a perfect crystal has not yet been theoretically considered ; no doubt the introduction of such statistical imperfections would profoundly modify the " perfect crystal " reflection formula. An attempt has already been made by Hall (1942) to estimate the elastic constants of KBr by means of observations of extra reflections, using Cu Kα radiation. Chemical purity and the absence of surface or other gross forms of distortion would be essential requirements for accurate and reproducible measurements.

When Mr. Smith and I first obtained extra-reflection photographs from crystals of organic compounds, the possibility that our diffuse maxima (especially in the case of benzil) might be similar to those observed for anthracene by Charlesby, Finch and Wilman was very much in our minds. As Drs. Charlesby and Wilman have pointed out, however, maxima of the electron-diffraction type would only be observable using very short x-ray wave-lengths, molecules of high internal periodicity and large-angle deflections. Even then they would be so diffuse as to cover very large areas of the background, and it is doubtful whether any really satisfactory observations could be made. It must be remembered that the reciprocal-lattice area covered by a single electron-diffraction photograph is many times larger than that explored by means of an x-ray photograph. I believe that the conditions of the different experiments also imply a further distinction between the methods. The use of a very thin film in the electron-diffraction method involves the extension of all reciprocal-lattice points along one direction in reciprocal space and the absence of long thermal waves in that particular direction. In effect the electron-diffraction method examines only two-dimensional thermal vibrations, the x-ray method three-dimensional ones. The two methods could perhaps best be made to overlap by a study of films of long-chain compounds or of substances of very high molecular weight and internal regularity. The Faxén-Waller theory could probably be extended to account for the diffuse electron maxima, but at present it does not include them.

The use of " extra " reflections in the examination of crystals of unknown structure has so far only been applied to fairly complicated organic compounds in which there is a chance that a layer or chain (or combined layer and chain, i.e. a lath-like) arrangement of molecules might occur. Either of these arrangements is most readily identified from a few well-exposed Laue photographs ; and the position of the layer or chain relative to the crystal axes can be unequivocally located. The method is not, of course, a substitute for the standard methods of analysis, but it has already given most useful supplementary, and even " leading ", information (Lonsdale, 1941; Lonsdale, Robertson and Woodward, 1941 ; Knaggs and Lonsdale, 1942).

The differences in the extra reflections from the two types of diamond cannot be explained as due to differences in the elastic constants, but, as Mr. Preston has remarked, any crystal imperfection, even if not spatially periodic, will produce a background effect, which may or may not be observable. The strength of the structure-sensitive secondary extra reflections from normal (type I) diamonds (and their temperature-insensitiveness) indicates a spatially periodic imperfection which in some specimens may attain considerable magnitudes.

No exact measurements of the size and intensity of the extra reflections over a range of intensity has yet been made for any substance for comparison with the predictions of the Faxén-Waller theory. An estimate of the expected effect according to the Bragg theory would be very difficult to make ; but if it is assumed that the effect of increasing temperature would be the breaking up of the structure into more and smaller crystallites, then, from a very general point of view, both theories would lead to the same expected result, an increase in both size and intensity of the extra reflections, particularly in those far from reciprocal-lattice points and of high order. I do not think that this could be made a crucial experiment, although the reversible changes with temperature which are experimentally found prove, of course, that no *temperature-independent* breaking up of the crystal into small groups can possibly explain the observed phenomena.

REFERENCES

Ewald, P. P. and Renninger, M., 1935. *Phys. Soc. Conference*, **2**, " The Solid State of Matter ", p. 60.

Fox, G. W. and Carr, P. H., 1931. *Phys. Rev.* **37**, 1622.

Fox, G. W. and Fraser, W. A., 1935. *Phys. Rev.* **47**, 899.

Fox, G. W. and Stebbins, D. W., 1939. *Phys. Rev.* **55**, 405.

Hall, J. H., 1942. *Phys. Rev.* **61**, 158.

Jauncey, G. E. M. and Bruce, W. A., 1938. *Phys. Rev.* **54**, 163.

Jauncey, G. E. M. and Deming, J. J., 1935. *Phys. Rev.* **48**, 462.

Jauncey, G. E. M. and Jaques, A. T., 1936. *Phys. Rev.* **50**, 672.

Klauer, F., 1935. *Phys. Z.* **36**, 208.

Knaggs, I. E. and Lonsdale, K. *J. Chem. Soc.* (in press).

Lonsdale, K., 1941. *Proc. Roy. Soc.* A, **177**, 272.

Lonsdale, K., Robertson, J. M. and Woodward, I., 1941. *Proc. Roy. Soc.* A, **178**, 43.

Nishikawa, S., Sakisaka, Y. and Sumoto, I., 1934. *Sci. Papers Inst. Phys. and Chem. Res., Tokyo*, **25**, 20.

Weigle, J. and Smith, C. S., 1942. *Phys. Rev.* **61**, 32.

THE SECONDARY X-RAY SPECTRUM OF SYLVINE: THEORY AND EXPERIMENT

BY (THE LATE) SIR WILLIAM BRAGG,
O.M., K.B.E., F.R.S.

MS. received 31 *May* 1942

NOTE BY SIR LAWRENCE BRAGG, F.R.S. AND MRS. KATHLEEN LONSDALE

This paper was written by the late Sir William Bragg shortly before his death, and was left in such a form that publication is possible, though we have ventured to make certain small alterations or additions where the original was incomplete, and have provided an abstract. The treatment of the extra reflections which appear at high temperatures is generally by means of the reciprocal lattice. Sir William Bragg expressed his results without making use of this analytical device, but in order to examine the correspondence between reciprocal and direct treatments he wrote a special section which appears in the paper. In the reciprocal treatment, the diffuse extra reflections, whose strength is enhanced by rise of temperature, can be represented as a cloud surrounding each point of the reciprocal lattice. This feature, equivalent to the construction in figure 1, follows equally either from the hypothesis that the spots are due to modulations of the lattice by thermal waves, or from the hypothesis that there are in the crystal small regions which are out of phase with each other in a completely random way. The main point made in the present paper is that the latter hypothesis, which leads to a simple optical treatment of the diffraction, is in better accord with the form of the observed spots in the case of KCl.

ABSTRACT. The co-existence in a crystal of very large and very small groups of atoms, the atoms within any group being in correct relative positions, but out of alignment with the atoms of neighbouring groups, so that all phase relationships between the scattered x-ray waves from successive groups is lost, would give both sharp Laue spots and diffuse background spots on a Laue photograph. This hypothesis predicts nearly circular diffuse spots for KCl at small angles of total deflection, whereas the Faxén-Waller theory, using known values of the elastic constants, predicts diffuse spots in the shape of elongated ellipsoids of maximum intensity at the centre. It is claimed that in this respect the hypothesis that the crystal is broken up into groups of atoms, with loss of continuity in phase relationship, is in better accord with observation than the hypothesis that it remains a single crystal distorted by the heat waves, but with no loss of continuity.

THE *extra pattern* in the diffraction of x rays by crystals has lately been explained as due to interaction between waves of the x rays and the thermal waves within the crystal. The theory on which the explanation is based has been discussed by Faxén, Waller, Laue, Ott, Zachariasen, Born and Sarginson, and Jahn. On account of the relations between the thermal movements and the elastic constants of the crystal, the calculated results can be expressed in terms of these elastic constants when they are known, and may then be compared with the experimental results.

In 1938 G. D. Preston made one of the first experimental examinations of the extra pattern. His crystals were NaCl, Al and MgO (Preston, 1939). He ascribes the effect, not directly to thermal movements, but to the presence of very small groups of atoms, each group a crystal in itself, imperfectly aligned with the main body of the crystal and therefore acting incoherently.* As will be shown presently—although this is not generally recognized—the hypothesis explains the extra pattern at least as well as the previous theory. It has, however, been objected to on various grounds, the principal being, in the first place, that it is unnecessary since the elastic-constant theory is sufficient, and, in the second place, that the presence of these small groups and their supposed mode of action would be difficult to reconcile with other crystal effects.

Nevertheless, it seems worth while to examine this theory, even if its fundamental assumptions are not warranted. It is remarkable that two theories which seem to be entirely different should lead to results which are very much alike, and are in more or less agreement with all experiments so far made. It may well be that though the two theories are formally different, they are actually linked closely together.

Let us then provisionally assume Preston's theory, and compare its predictions with experiments.

A very small group of atoms scatters an incident beam of x rays in a more diffuse manner than a crystal of larger size. In fact the diffuseness is often used as an indication of the size of the group.

Most of the experiments which lend themselves to the comparison between these theories and experiment have been made with crystals which are actually monatomic, such as diamond and sodium, or effectively so, such as sylvine. The crystal structure of the latter is formed by the repetition of the atom at constant intervals in three directions, so related that three single displacements, one in each direction, form three edges of a unit cell: that is to say, a parallelopiped having atoms at its corners only, and none within it or on its faces.

Suppose that a group is a larger parallelopiped whose edges are parallel to the three chosen directions, and that there are p atoms on one edge, inclusive of those at each end, q on another, and r on the third; then the amplitude of the scattered ray is proportional to

$$\frac{\sin p\phi_a/2}{\sin \phi_a/2}\cdot\frac{\sin q\phi_b/2}{\sin \phi_b/2}\cdot\frac{\sin r\phi_c/2}{\sin \phi_c/2},$$

where ϕ_a, ϕ_b, ϕ_c are the differences in phase between the rays scattered at that corner of the cell which has been chosen as origin and the three neighbouring corners respectively.

---

* *Preston, p.* 122, *loc. cit.* "The fact that the intensity of the spots increases as the temperature of the crystal is raised suggests that their presence is connected with the temperature vibrations of the lattice . . . The effect of this wave-like disturbance of the regularity of the lattice is to produce a series of "ghosts" associated with the normal spectra; if the disturbance consists of a very large number of lattice waves these will produce a continuous faint background on the Laue photographs. The experiments described above show that this background is not of uniform intensity but possesses well-marked maxima. The simplest explanation of the presence of these maxima (in aluminium single crystals) is to suppose that the thermal vibrations are of such a nature that the crystal is divided up into small groups of atoms; in each group the face-centred arrangement is accurately maintained, but neighbouring groups have slightly different spacings due to a wave of dilation, or are inclined at a small angle to the average direction of the crystal as a result of a wave of shear." [*Note added by Sir W. L. Bragg and Mrs. K. Lonsdale.*]

If the direction cosines of the scattered and incident rays, referred to rectangular axes having the chosen corner as origin, be h, k, l, h', k', l' respectively, if x_1, y_1, z_1 be the coordinates of the atom at the end of the first repetition along the first axis, and so on, and if λ be the wave-length, then

$$\phi_a = \frac{2\pi}{\lambda}\{(h-h')x_1+(k-k')y_1+(l-l')z_1\}.$$

In the case of a group consisting of a single cell, there are two atoms on each edge, and the expression reduces to

$$8\cos\phi_a/2 \,.\, \cos\phi_b/2 \,.\, \cos\phi_c/2.$$

Whatever the value of p, q, r, the expression is at a maximum when ϕ_a, ϕ_b, ϕ_c are all multiples of 2π. In this case, the corners of the unit cell, and the atoms throughout the group, are in phase with each other, as in the case of a Bragg reflection. There are subsidiary maxima when only two of the three differences of phase are multiples of 2π: and others of still less importance when only one of them is such a multiple.

In any given case when the relative coordinates of the cell corners and also the direction of the incident ray are known, equations may be formed which are the condition for the realization of the maxima. In practice it is convenient to let the incident rays be parallel to one axis: thus its direction cosines may be taken to be 001. Then the maxima occur when

$$hx_1+ky_1+(l-1)z_1=M_1\lambda,$$
$$hx_2+ky_2+(l-1)z_2=M_2\lambda,$$
$$hx_3+ky_3+(l-1)z_3=M_3\lambda,$$

where M_1, M_2, M_3 are any integers, λ is the wave-length, and x_1, y_1, z_1; x_2, y_2, z_2; x_3, y_3, z_3 are the coordinates of the three corners of the cell nearest to the origin. These equations give the values of h, k, l for which the corner of the cell chosen as origin are in phase with the three neighbouring corners respectively. They may be looked on as the equations in h, k, l of planes cutting the sphere $h^2+k^2+l^2=1$.

Choosing any two of the three direction cosines, such as h, k, curves may be plotted which represent the projection of the intersections of planes and sphere upon a plane perpendicular to the incident ray. In practice, this plane generally coincides with that of the photographic film or plate, so that there is a direct and simple relation between h, k and the position of the point on the plate where the scattered ray strikes it. If ξ, η are the coordinates of that point referred to axes parallel to the main axes, but having as origin the point where the direction of the incident ray strikes the plate, then $\xi = hD/l$, $\eta = kD/l$, where D is the distance from the crystal to the plate.

The adjoining figure 1, which is reproduced from an article in the *Proceedings of the Royal Society* (W. H. Bragg, 1941), illustrates these points. The crystal is sylvine, which may be considered to be a simple cubic crystal of side 3·14 A. The incident Mo rays, for which $\lambda = 0{\cdot}71$ A., make an angle of 20° with one face of the cube and are parallel to an adjoining face. The coordinates of the three corners referred to the fourth, as described above, are then $(a\cos\theta, 0, -a\sin\theta)$: $(0, a, 0)$: $(a\sin\theta, 0, a\cos\theta)$, where $a = 3{\cdot}14$, $\theta = 20°$.

The equations become

$$\begin{aligned} a\{h\cos\theta-(l-1)\sin\theta\} &= M_1\lambda, \\ ak \qquad\qquad &= M_2\lambda, \\ a\{h\sin\theta+(l-1)\cos\theta\} &= M_3\lambda. \end{aligned}$$

When arithmetical values are given to the quantities a, θ and λ, three sets of (h, k) curves for integral values of M_1, M_2, M_3 can be calculated. The (ξ, η) curves are readily derived, by means of the relations given above, and are shown in the figure, the scale being chosen so that D, the distance from crystal to plate, has the value 10. The first equation gives the curves which are approximately vertical; the second, those which are approximately horizontal; and the third the three more widely spaced curves, of which one is almost a semicircle. Only one-half of the complete diagram is shown, the second half being the image of the first.

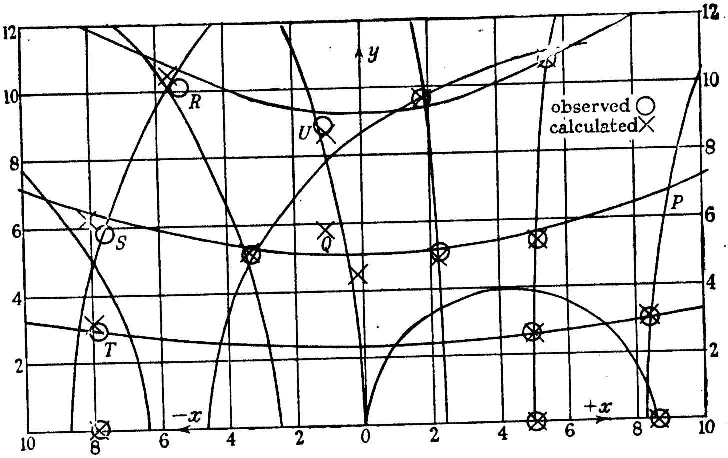


Figure 1.

Although there is in this diagram no triple point of intersection of members of the three sets of curves, there are several places where three curves nearly meet in a point. When that occurs, the corners of the unit cell are nearly in phase with each other and an observable scattering becomes a possibility. If members of the three sets met in a single point, there would be a Bragg reflection for the particular direction (h, k) given by the coordinates of that point. The small round circles show where the maxima of the extra spots under discussion are observed in the photograph; it will be observed that their positions coincide with expectation to a remarkable degree. Spots are found where members of the first two sets of curves intersect, provided a member of the third set is not too far away.

For the purpose of a closer comparison between theory and experiment, the positions of the maxima have been calculated for the case when $p = q = r = 2$, i.e. the amplitude has been put proportional to $\cos\phi_a/2 \,.\, \cos\phi_b/2 \,.\, \cos\phi_c/2$. Whatever the actual values of p, q and r, the positions of the principal maxima are not sensibly affected: groups of more than the eight atoms at the corners of a cubic cell have secondary maxima in various positions, and these add a diffused background which diminishes the contrast between the primary maxima and their surroundings. The calculated values are marked by crosses, and their close agreement with the small circles shows that the theory gives in this case a satisfactory explanation of what is observed. Theory presumes the existence of small groups of no particular shape in slight disarray with the main crystal and with one another. The slight disarray does not prevent them from acting simultaneously. If the crystal is so set that the main body is giving a Bragg reflection, the action of the groups will merely cause a little diffused effect on either side of the main reflection, which, moreover, will absorb most of the energy, and the diffused effect will consequently be weaker.

Suppose that there was a small group alongside a large one, ranged more or less parallel to it. If, then, a beam of monochromatic rays fell upon the two, and the inclination of the pair to the beam were gradually altered, over a range which covered the point where the incident rays experienced a Bragg reflection in the large crystal, then the reflection from the large group would flash out at the moment when the inclination of the rays to the crystal planes was right, but would vanish at all other angles. On the other hand, there would be a reflection—much feebler, of course—from the small group which would persist over a certain angle which might be considerable, say 10° or 20°. At the same time the inclination of the scattered ray to the incident would not vary very much, as generally happens in the case of diffraction gratings, and is found to be the case with these extra spots. The extra spot would not change its position much on the plate during the turning of the crystal; the Bragg reflection would flash out for an instant, and would at that instant coincide with the other spot.

If the radiation contained rays of other wave-lengths (including that of the strong monochromatic ray), there would be no change in the extra-spot phenomenon, because only the strong monochromatic beam could produce it. But there would be a reflection from the large group of atoms at all times; whatever the position of the crystal, the white radiation would include some wave-length which would be reflected by the large group. During the turning, the resulting spot would move towards the extra spot and go through it. The extra spot would brighten at the approach of the reflection from the large group, because when the inclination is right for reflection from the large group it is at least nearly right for reflection from the small group, the planes of the two being at least approximately parallel.

The form of the extra spot is nearly circular in most cases, certainly when the spot, as in the photograph corresponding to figure 1, is one of the dozen or so spots near the origin. Calculations based on the formulae already given agree with this experimental result: it is, in fact, to be expected from the simple causes of its origin as given by the theory at present under consideration. The small groups are supposed to consist of atoms few in number in each case; and in each case the atoms are arranged among themselves in the manner of the particular crystal to which they belong. They may be of any form provided that the average form is spherical, that is to say, there is no preponderance of the dimensions of the groups in particular directions.

In the foregoing, an immediate comparison is made between theory and experiment by putting together the sylvine photograph and the corresponding (h, k) diagram with its immediate derivative, the (ξ, η) diagram.

In 1938 and 1939 Laval described a long and careful series of measurements of the background x-ray diffraction, which is equivalent to saying that he investigated the extra-spot effect. He used an ionization spectrometer: keeping the spectrometer fixed, and, therefore, confining attention to a particular value of (h, k), he examined the result for different positions of the crystal. In the photographic method the crystal is fixed, and all values of h and k are considered simultaneously. It is not a simple matter to compare in detail the results obtained by the two methods, but the general result is quite clear. Laval's experiments agree, so far as can be seen, with those of the photograph, as must be the case because the subject is the same. Laval's results are expressed in terms of a reciprocal lattice, and it is convenient therefore to use this lattice in an examination of his results and a comparison of them with the (h, k) diagram. Expression of results in terms of the (h, k) diagram and expression in terms of the reciprocal lattice can be linked together in the following simple way.

Suppose that in figure 2, IC represents the direction of a bundle of monochromatic x rays incident on a crystal at C. Let this direction be the axis of z, and let the axis of x be in the plane of the paper, as shown, and the axis of y be perpendicular to that plane.

Let (h, k, l) be the direction cosines of any line CQ which does not necessarily lie in the plane of the paper.

Let RR′ and CP be the external and internal bisectors of the angle ICQ lying in the plane ICQ. Let (l', m, n) be the direction cosines of CP. Then $h = -2l'n$; $k = -2mn$; $l = 1 - 2n^2$ and $n = \sin \mathrm{ICR}'$.

Let $\mathrm{CP} = 2K^2n/\lambda$, where K is a constant for different values of h, k. The locus of P is a sphere with centre O and radius OC, as shown in the figure, because if PI′ is perpendicular to CP, then $\mathrm{CI}' = \mathrm{CP}/n = 2K^2/\lambda$ (O is the mid-point of CI′).

Let A be any point; let $\mathrm{CA} = a$; and let (p, q, r) be the direction cosines of CA. If rays parallel to (001) are incident on C and A, and are scattered in both cases in the direction (h, k, l), the difference in length of path of the two rays is

$$pah + qak + ra(l-1),$$

since x, y, z for A will be pa, qa, ra.

If the scattered rays are in phase, this quantity is an integral multiple of λ. Then

$$\begin{aligned} M\lambda &= pah + qak + ra(l-1) = -2pal'n - 2qamn - 2ran^2 \\ &= -2na(pl' + qm + rn) = -2na \cos \mathrm{PCA} \\ &= -\frac{\lambda a}{K^2} . \mathrm{CP} \cos \mathrm{PCA} \quad \left(\text{since } 2n = \frac{\lambda}{K^2} . \mathrm{CP}\right) \\ &= -\frac{\lambda a}{K^2} . \mathrm{CA}', \text{ where CA}' \text{ is the projection of CP on CA.} \end{aligned}$$

Hence $\mathrm{CA}' = M . K^2/a$, where M is any integer.

If, therefore, the values of h, k are such that C and A are in phase, P lies on a plane passing through A′ and perpendicular to CA: this is the reciprocal plane of the point A.

As a result of the definition, P must also lie on the sphere of radius K^2/λ, centre O, and has no significance unless it does so. This is really equivalent to the statement that the values of h, k, l are such that $h^2+k^2+l^2=1$.

If one corner of the cubic cell of sylvine be taken as origin, an adjacent corner is in phase for a given value of h, k if the point P corresponding to $(h\ k)$ lies on one of a set of parallel planes spaced at intervals of K^2/a, one of them $(M=0)$ going through the origin. In practice it is convenient to make the radius of the sphere K^2/λ equal to 10 units. The spacing is therefore $10\lambda/a$; when Mo rays are used, this is equal to 2·26.

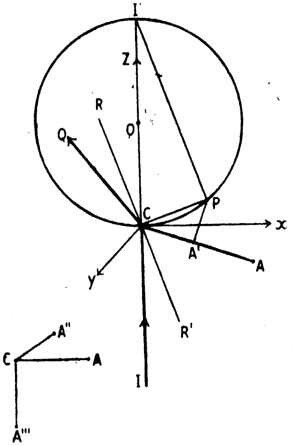


Figure 2. Q, RR′, P, A and A′ are not necessarily in the plane of the paper, although they have been drawn in that way for convenience.

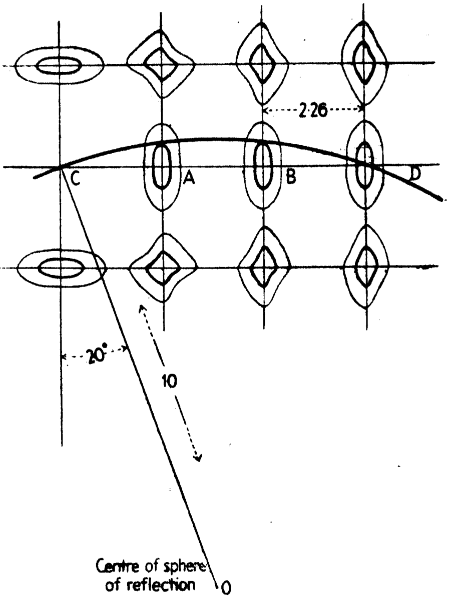


Figure 3. Isodiffusion surfaces and part of sphere of reflection for KCl. $\theta_{100}=20^\circ$. MoKα radiation; as calculated from Faxén theory.

The triple series of planes corresponding to the three cube corners A, A″, A‴ forms a set of cubic boxes. If the point P lies on a circle of intersection of any of these planes by the sphere, the corresponding values of (h, k) are those for which the origin and one other of the cube corners are in phase. If it lies at a point where the sphere cuts a line of intersection of two planes, then for that value of (h, k) three corners of the unit cell are in phase with each other. If the relative orientations of the incident ray and the crystal happen to be such that the sphere passes through a triple point of intersection, then for the corresponding value of (h, k) there is a Bragg reflection of the monochromatic ray.

With each position of P may be associated a quantity proportional to the intensity of the ray scattered in the corresponding (h, k) direction, which can, of course, only have an existence when sphere and crystal are so adjusted that the sphere passes through that position. At a triple point, which Laval calls a node, the intensity is a maximum and falls away if P moves away from that point in any direction. Laval has mapped out this field round a nodal point of sylvine and found it to be of a globular form, rather flattened on one side. This result may be compared with that of the photographic method in the following way:—

When P is projected upon the *xy* plane, its coordinates are l' . CP and m . CP, which are equal to $2K^2l'n/\lambda$ and $2K^2mn/\lambda$, or $-K^2h/\lambda$, $-K^2k/\lambda$.

When K^2/λ is put equal to 10, then these coordinates are equal to $-10h$ and $-10k$ respectively.

Thus the intensity associated with any position of P when projected on the *xy* plane is the same as that corresponding to the point (h, k) in the (h, k) diagram. As this latter is very nearly a true map of the intensities of the photograph, at least when l is nearly equal to 1, so also is the map formed by the projection of all the points P, where the sphere cuts the field round the node, the corresponding intensity defined by the field being associated with the projection at each point. A test of this construction shows that the spots are approximately circular when the spots are fairly close to the origin: and the intensity field round the node must be such that the projection is of circular form, with a maximum at the centre and an equal fading away in all directions. Laval's results satisfy the test fairly well.

This condition is, however, far from being satisfied if the field round a node is as calculated on the theory of Faxén and subsequently developed by several other writers. Mrs. Lonsdale has been so kind as to draw for me maps of sections of the reciprocal lattice, from which figure 3 is an extract. The plane of the section is perpendicular to the y axis and goes through C. The incident rays are in the plane of the paper, and are, as before, taken to define the axis of z; and one side of the squares is inclined at 20° to the axis of z. Calculation based on the Faxén theory indicates that the reflecting regions or " fields " round ABC are oblate solids of revolution, in which the small axis along the edge of the cube CABD is less than one-half of the longer axes in the other principal directions. The section of these solids of revolution by the sphere in the neighbourhood of A, B and D is in each case nearly of elliptical form.

The projection of the section upon a plane perpendicular to the axis of z gives the form of the corresponding spot in the photograph, which thus, according to the Faxén theory, should be elongated. On the other hand, the small-group hypothesis gives round spots, and this, it is claimed, corresponds more nearly to the form actually observed both by Laval's measurements and in photographs.

REFERENCES

Bragg, Sir W. H., 1941. *Proc. Roy. Soc.* A, **179,** 54.

Laval, J., 1938. *C.R. Acad. Sci., Paris,* **207,** 169 ; 1939. *Ibid.* **208,** 1512 ; *Bull. Soc. Franç. Min.* **62,** 137.

Preston, G. D., 1939. *Proc. Roy. Soc.* A, **172,** 116.

LATTICE DYNAMICS AND X-RAY SCATTERING

By MAX BORN, F.R.S.,
The University, Edinburgh

Communicated by Thos. J. Lonsdale, 3 March 1942

ABSTRACT. The " extra spots " on Laue photographs are considered by many investigators in Europe and America as due to the thermal motion of the crystal lattice, but a group of Indian physicists led by Raman have proposed a different theory which they claim to be based on the assumption of a new specific quantum effect. Raman criticizes in particular the current theory of lattice vibrations as it makes use of the *method of the cyclic lattice*, which he considers wrong. The present paper refutes these criticisms. It is shown that the cyclic solution is in complete agreement with the rigorous solution in the case of a diatomic linear lattice where an optical branch exists in addition to the acoustical one. Raman's theory is compared with the dynamical lattice theory, which can be developed on rigorous quantum methods, while Raman's considerations are essentially classical. His main error consists in considering the velocity of wave propagation in a lattice as constant ; in fact, it is a function of the wave number, which for the acoustical branches varies in such a way that the intensity of scattering has maxima in the neighbourhood of the Laue spots. Hence the extra spots are due to the acoustical branches, not the optical ones, as Raman contends. The paper further contains a short account of the results of the dynamical theory obtained by Zachariasen, Jahn, and Sarginson, which are in good agreement with the experiments of Lonsdale and Smith, Zachariasen and others.

§ 1. INTRODUCTION

THE theory of the new type of spots on Laue photographs ("extra spots") is still in a controversial state. The thermal theory * (Debye, 1913 and 14; Faxén, 1923; Waller, 1923, 1928), which is accepted by Lonsdale and Smith (1941), Zachariasen (1940, 1941) and other investigators, is opposed by Sir C. V. Raman (1942), who claims that the extra spots are due to a peculiar new quantum effect. His ideas are described in detail in a special number of *Proc. Ind. Acad. of Science*, which is completely devoted to papers on the new phenomenon. In order to justify his assumptions, Raman rejects some well-established results of the general theory of lattice dynamics, namely, the statement that the elastic spectrum consists of a set of branches or bands, each containing *N* frequencies (where *N* is the number of cells), and that the density of the frequencies in each band is uniform in the space of wave numbers. Raman declares that this distribution applies only to the low-frequency (or acoustical) branch which is used for the calculation of the specific heat (Debye's formula), whereas instead of the high-frequency (optical) branches there are sharply monochromatic vibrations. He believes that he has strong experimental evidence for this view. I shall not enter into the question of the experimental confirmation of the different theories, as Mrs. Lonsdale, in her lecture, has done this much more expertly. I shall restrict myself to the discussion of the theoretical side of the problem. Raman also puts forward theoretical arguments for his view; he considers as the weak spot of the lattice dynamics the assumption of the *cyclic lattice*. I quote his words: "The so-called postulate of the 'cyclic lattice' is in fact nothing more than an *ad hoc* hypothesis adopted for mathematical convenience", and he adds a critical examination of it, which culminates in the assertion that the usual method of Fourier analysis of the atomic vibrations "would result in the phases of the individual oscillators being entirely arbitrary and uncorrelated with each other. This is evidently *a reductio ad absurdum*, since such a state of affairs could only exist when the Planck oscillators are entirely independent of each other, as would be the case if we were considering the molecules of a gas. It is a fundamental aspect of the crystalline state that the oscillators of which it is composed are coupled to each other more or less firmly". The last remark seems to me due to a misunderstanding. In lattice dynamics the atoms are, of course, supposed to be coupled to each other. Just because they *are* coupled, they do not represent Planck oscillators to which quantization could be applied. But if the forces are linear in the displacements, one can always introduce normal coordinates which are independent oscillators and can be quantized. They are used in the statistical mechanics of crystals (specific heat, thermal expansion, etc.), and it is obvious that they also provide the method of treating the influence of lattice vibrations on x-ray scattering (Debye 1913–14).

---

* My own work, partly in collaboration with Miss K. Sarginson (the first part of which appeared last year), is a modernized and more rigorous form of Waller's theory.

More important is Raman's critical attitude to the method of the cyclic lattice; for it is perfectly correct that it was invented (Born, 1923) for mathematical convenience, and that there is no proof apart from an almost trivial case. The method consists in replacing the real boundary conditions of a vibrating lattice by simplified ones, namely, the conditions of periodicity in a large parallelopipedic domain. There are several physical arguments in favour of this procedure: the real surface is unknown; it is very likely not smooth, but quite irregular and rough. We are not interested in properties of the surface, or in properties depending on the surface, but only in intrinsic properties of the matter forming the crystal. Further, mathematicians (Weyl, 1911, 12, 14, 15, and Courant, 1920) have proved that in a continuous medium the asymptotic distribution of vibrations for high frequencies does not depend on the shape of the surface nor on the boundary conditions. It is likely that the same should hold for a vibrating discontinuum. From the standpoint of physics, the cyclic condition is not without experimental confirmation. It is used in the electronic theory of metals, where it leads to a distribution of bands of electronic energy, called Brillouin zones, which are very similar to the frequency bands of the lattice vibrations. This band structure of electronic levels is the basis of the modern theory of mechanical and electrical properties of metals; it explains, for instance, the different types of conductivity in their relation to the periodic table of elements, and many other facts.

But I quite agree with Raman that the cyclic method is not proved, apart from the case of the linear monatomic lattice. Here the situation is this:—Let a chain of N equal particles (mass m) at equal distances a be coupled by quasi-elastic forces acting between neighbours only, equal to α per unit displacement; the ends may be either free or coupled to two fixed points at a distance a from the ends by the same force α. Let the displacement of particle l for the normal mode j of frequency ω_j be u_l; then, omitting the time factor $\cos(\omega_j t+\epsilon_j)$, we have (Born and Karman, 1912):

$$\left.\begin{array}{ll} \text{Free ends:} & \\ u_l=u\cos(l-\tfrac{1}{2})\phi_j, \quad \omega_j=2\sqrt{\left(\dfrac{a}{m}\right)}\left|\sin\dfrac{\phi_j}{2}\right|, \quad \phi_j=\dfrac{\pi j}{N} & \\ \text{Fixed ends:} & \\ u_l=u\sin l\phi_j, \quad \omega_j=2\sqrt{\left(\dfrac{a}{m}\right)}\left|\sin\dfrac{\phi_j}{2}\right|, \quad \phi_j=\dfrac{\pi j}{N+1} & \end{array}\right\} j=1,\,2,\ldots.N.$$

This rigorous solution may be compared with that obtained with the cyclic condition, which is

$$u_l=ue^{il\phi_j}, \quad \omega_j=2\sqrt{\left(\frac{a}{m}\right)}\left|\sin\frac{\phi_j}{2}\right|, \quad \phi_j=\frac{2\pi j}{N}, \quad j=1,2,\ldots.\frac{N}{2}.$$

In all cases the distribution of the frequencies over the scale of wave number (j-scale) is uniform; but in the total range of frequencies $\left(0<\omega_j\leqslant 2\sqrt{\left(\frac{\alpha}{m}\right)}\right)$ there are apparently twice as many frequencies in the rigorous solution as in the cyclic solution. The reason is simply that in the latter case each solution is complex, hence represents really two solutions (corresponding to two independent running waves with opposite equal velocities), while the correct solution is one standing wave (or two opposite running waves with constant phase relation). The total number of frequencies is therefore exactly the same in both cases. For small N the cyclic spectrum is rather different from the real one; but for large N both tend to the same uniform density of frequencies.

This example shows only one branch and does not help to clarify the situation in the higher branches with which Raman's theory is concerned. I propose, therefore, to treat here another simple example, the di-atomic linear lattice. I shall show in the following section that in this case, also, the cyclic method leads, for large N, to the same result for both branches as the correct solution for the finite chain.

The case of a 3-dimensional lattice brings in a new feature, the shape of the boundary. Here a simple result can only be expected asymptotically for high wave numbers, as in the case of continous media.

Assuming that the uniform density of vibrations in the space of wave numbers (reciprocal space) is correct, there is no ambiguity in the theory of x-ray scattering. I shall, in the last section, give a short account of this theory and show where Raman's considerations * are inaccurate or incorrect.

§ 2. THE LINEAR DIATOMIC CRYSTAL

Let N atoms of mass m_1 and N atoms of mass m_2 (figure 1) form an equidistant alternating chain, which is an equilibrium configuration if all forces except those

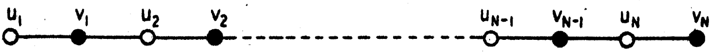


Figure 1.

between first neighbours are negligible. If u_1, u_2, . . . u_N are the (longitudinal)

---

* I shall not use Raman's notation, but write his formulae in such a way that they can be readily compared with the general theory.

displacements of the first kind of atoms, $v_1, v_2, \dots v_N$ those of the second kind, the equations of motion for small vibrations are

$$\begin{array}{ll}
m_1\ddot{u}_1=\alpha(v_1-u_1), & m_2\ddot{v}_1=\alpha(u_2-v_1)-\alpha(v_1-u_1),\\
m_1\ddot{u}_2=\alpha(v_2-u_2)-\alpha(u_2-v_1), & m_2\ddot{v}_2=\alpha(u_3-v_2)-\alpha(v_2-u_2),\\
\cdots\cdots & \cdots\cdots\\
m_1\ddot{u}_l=\alpha(v_l-u_l)-\alpha(u_l-v_{l-1}), & m_2\ddot{v}_l=\alpha(u_{l+1}-v_l)-\alpha(v_l-u_l),\\
\cdots\cdots & \cdots\cdots\\
m_1\ddot{u}_N=\alpha(v_N-u_N)-\alpha(u_N-v_{N-1}), & m_2\ddot{v}_N=\qquad\qquad -\alpha(v_N-u_N).
\end{array} \qquad \dots\dots(1)$$

The difficulty of this problem lies in the deviation of the first and the last equation from the general form of all the rest; this deviation is the expression of the correct boundary conditions, namely, that the first and last particles are subjected to no external forces.

The cyclic device overcomes this difficulty by adding the missing terms, writing for the first and last equations

$$\begin{array}{ll}
m_1\ddot{u}_1=\alpha(v_1-u_1)-\alpha(u_1-v_0), & \cdots\cdots\\
\cdots\cdots & m_2\ddot{v}_N=\alpha(u_{N+1}-v_N)-\alpha(v_N-u_N),
\end{array}$$

and adding the condition that $v_0=v_N$, $u_{N+1}=u_1$. Then the finite chain can be extended to infinity on both sides by periodic repetition, so that the general conditions $u_l=u_{N+l}$, $v_l=v_{N+l}$ (which contain the two special ones) hold. The modified problem can be immediately solved.

I shall, however, give a rigorous solution of the equations (1). We consider a harmonic vibration with frequency ω, where all the u_l, v_l are proportional to $\exp(-i\omega t)$. With the abbreviations

$$\tfrac{1}{2}(m_1+m_2)=m, \qquad m\omega^2/2\alpha=\epsilon, \qquad \dots\dots(2)$$

the equations can be written

$$\begin{array}{ll}
\left(1-2\frac{m_1}{m}\epsilon\right)u_1=v_1, & \left(1-\frac{m_2}{m}\epsilon\right)v_1=\tfrac{1}{2}(u_1+u_2),\\
\left(1-\frac{m_1}{m}\epsilon\right)u_2=\tfrac{1}{2}(v_1+v_2), & \left(1-\frac{m_2}{m}\epsilon\right)v_2=\tfrac{1}{2}(u_2+u_3),\\
\cdots\cdots & \cdots\cdots\\
\left(1-\frac{m_1}{m}\epsilon\right)u_l=\tfrac{1}{2}(v_{l-1}+v_l), & \left(1-\frac{m_2}{m}\epsilon\right)v_l=\tfrac{1}{2}(u_l+u_{l+1}),\\
\cdots\cdots & \cdots\cdots\\
\left(1-\frac{m_1}{m}\epsilon\right)u_N=\tfrac{1}{2}(v_{N-1}+v_N), & \left(1-2\frac{m_2}{m}\right)v_N=u_N.
\end{array} \qquad \dots\dots(3)$$

We try to solve these equations by expressions of the form (omitting the time factor)

$$\begin{aligned}
u_l&=A_1\cos l\phi+A_2\sin l\phi,\\
v_l&=B_1\cos l\phi+B_2\sin l\phi,
\end{aligned} \qquad \dots\dots(4)$$

and substitute these first in the two equations with index l. We easily find

$$\begin{aligned}
\tfrac{1}{2}(u_l+u_{l+1})&=\cos\phi/2\{\cos l\phi(A_1\cos\phi/2+A_2\sin\phi/2)\\
&\qquad+\sin l\phi(-A_1\sin\phi/2+A_2\cos\phi/2)\},\\
\tfrac{1}{2}(v_{l-1}+v_l)&=\cos\phi/2\{\cos l\phi(B_1\cos\phi/2-B_2\sin\phi/2)\\
&\qquad+\sin l\phi(B_1\sin\phi/2+B_2\cos\phi/2)\}.
\end{aligned}$$

By equating the factors of $\cos l\phi$ and $\sin l\phi$ on both sides of these equations, four conditions are obtained:

$$\begin{aligned}
\left(1-\frac{m_1}{m}\epsilon\right)A_1&=\cos\phi/2(B_1\cos\phi/2-B_2\sin\phi/2),\\
\left(1-\frac{m_1}{m}\epsilon\right)A_2&=\cos\phi/2(B_1\sin\phi/2+B_2\cos\phi/2),\\
\left(1-\frac{m_2}{m}\epsilon\right)B_1&=\cos\phi/2(A_1\cos\phi/2+A_2\sin\phi/2),\\
\left(1-\frac{m_2}{m}\epsilon\right)B_2&=\cos\phi/2(-A_1\sin\phi/2+A_2\cos\phi/2).
\end{aligned}\qquad\ldots\ldots(5)$$

These are 4 linear equations for A_1, A_2, B_1, B_2. The determinant of the 4th order is the square of a determinant of the 2nd order; each root is double, and therefore 2 of the 4 quantities A_1, A_2, B_1, B_2 can be chosen arbitrarily. One sees this easily by substituting, for instance, B_1, B_2 from the 3rd and 4th equation in the first two, which reduce to

$$\begin{aligned}
\left(1-\frac{m_1}{m}\epsilon\right)\left(1-\frac{m_2}{m}\epsilon\right)A_1&=A_1\cos^2\phi/2,\\
\left(1-\frac{m_2}{m}\epsilon\right)\left(1-\frac{m_1}{m}\epsilon\right)A_2&=A_2\cos^2\phi/2.
\end{aligned}$$

Hence A_1, A_2 are indeterminate, and the 2nd order determinantal equation is

$$\left(1-\frac{m_1}{m}\epsilon\right)\left(1-\frac{m_2}{m}\epsilon\right)-\cos^2\phi/2=0;\qquad\ldots\ldots(6)$$

the solution of this quadratic equation is

$$\left.\begin{matrix}\epsilon^{(1)}\\ \epsilon^{(2)}\end{matrix}\right\}=\frac{m^2}{m_1m_2}\left[1\mp\sqrt{\left(1-\frac{m_1m_2}{m^2}\sin^2\phi/2\right)}\right],\qquad\ldots\ldots(7)$$

or, if we put

$$\begin{aligned}
m_1&=m(1+\mu), & \mu&=\tfrac{1}{2}\frac{m_1-m_2}{m_1+m_2},\\
m_2&=m(1-\mu), & \frac{m_1m_2}{m^2}&=1-\mu^2,
\end{aligned}\qquad\ldots\ldots(8)$$

$$\left.\begin{matrix}\epsilon^{(1)}\\ \epsilon^{(2)}\end{matrix}\right\}=\frac{1}{1-\mu^2}[1\mp\sqrt{(\cos^2\phi/2+\mu^2\sin^2\phi/2)}].\qquad\ldots\ldots(9)$$

We now substitute the expressions (4) into the first and last equation of the set (3); this gives

$$\begin{aligned}
\left(1-2\frac{m_1}{m}\epsilon\right)(A_1\cos\phi+A_2\sin\phi)&=B_1\cos\phi+B_2\sin\phi,\\
\left(1-2\frac{m_2}{m}\epsilon\right)(B_1\cos N\phi+B_2\sin N\phi)&=A_1\cos N\phi+A_2\sin N\phi.
\end{aligned}\qquad\ldots\ldots(10)$$

Here we eliminate two of the constants, say A_1 and A_2, with the help of (5); then we obtain

$$\left(1-2\frac{m_1}{m}\epsilon\right)\left(1-\frac{m_2}{m}\epsilon\right)\frac{1}{\cos\phi/2}(B_1\cos\phi/2+B_2\sin\phi/2)=B_1\cos\phi+B_2\sin\phi,$$

$$\left(1-2\frac{m_2}{m}\epsilon\right)(B_1\cos N\phi+B_2\sin N\phi)$$
$$=\left(1-\frac{m_2}{m}\epsilon\right)\frac{1}{\cos\phi/2}[B_1\cos(N-\tfrac{1}{2})\phi+B_2\sin(N-\tfrac{1}{2})\phi]. \qquad \ldots\ldots(11)$$

The first of these equations can be written

$$B_1\left\{\left(1-2\frac{m_1}{m}\epsilon\right)\left(1-\frac{m_2}{m}\epsilon\right)-\cos\phi\right\}$$
$$+B_2\left\{\left(1-2\frac{m_1}{m}\epsilon\right)\left(1-\frac{m_2}{m}\epsilon\right)\tan\phi/2-\sin\phi\right\}=0,$$

and with the help of (6) simplified to

$$B_1=B_2\frac{1-\frac{m_2}{m}\epsilon}{\frac{m_2}{m}\epsilon}\tan\phi/2. \qquad \ldots\ldots(12)$$

The second equation (11) can be written

$$\cos N\phi\left\{B_1\frac{m_2}{m}\epsilon-B_2\left(1-\frac{m_2}{m}\epsilon\right)\tan\phi/2\right\}$$
$$+\sin N\phi\left\{B_2\frac{m_2}{m}\epsilon+B_1\left(1-\frac{m_2}{m}\epsilon\right)\tan\phi/2\right\}=0.$$

Here the coefficient of $\cos N\phi$ vanishes in virtue of (12), and that of $\sin N\phi$, namely (using (12) again)

$$B_2\left\{\left(\frac{m_2}{m}\epsilon\right)^2+\left(1-\frac{m_2}{m}\epsilon\right)^2\tan^2\phi/2\right\}\Big/\frac{m_1}{m}\epsilon,$$

does not vanish identically. Hence

$$\sin N\phi=0, \qquad \phi_j=\frac{\pi}{N}j, \qquad j=1, 2, \ldots\ldots N. \qquad \ldots\ldots(13)$$

The range of the values of j is determined by the consideration that, according to (9), the range of ϕ for which the frequency assumes each possible value just once, is given by $0<\phi\leqslant\pi$.

This proves that the distribution of frequencies is the same (for each branch $\epsilon^{(1)}$, $\epsilon^{(2)}$) as in the case of a monatomic chain, namely uniform in the j-scale.

This rigorous solution may now be compared with that obtained from the cyclic conditions (Born and Goeppert-Mayer, 1933). One has in this case the complex solution

$$u_l=ue^{il\phi}, \qquad v_l=ve^{il\phi}; \qquad \ldots\ldots(4')$$

the amplitudes u, v have to satisfy

$$u\left(1-\frac{m_1}{m}\epsilon\right)-v\tfrac{1}{2}(1+e^{-i\phi})=0,$$
$$-u\tfrac{1}{2}(1+e^{i\phi})+v\left(1-\frac{m_2}{m}\epsilon\right)=0; \qquad \ldots\ldots(5')$$

the matrix is Hermitean, but the determinantal equation is the same as before, (6), namely,

$$\begin{vmatrix} 1-\dfrac{m_1}{m}\epsilon & -\tfrac{1}{2}(1+e^{-i\phi}) \\ -\tfrac{1}{2}(1+e^{i\phi}) & 1-\dfrac{m_2}{m}\epsilon \end{vmatrix} = \left(1-\frac{m_1}{m}\epsilon\right)\left(1-\frac{m_2}{m}\epsilon\right) - \cos^2\phi/2 = 0, \qquad \ldots\ldots(6')$$

and, therefore, again we reach the expression (9) for $\epsilon^{(1)}$, $\epsilon^{(2)}$.

The set of admissible values for ϕ is obtained from the condition of periodicity, which leads to

$$\phi_j = \frac{2\pi}{N}j, \qquad j=1, 2, \ldots \frac{N}{2}. \qquad \ldots\ldots(13')$$

This differs from (13), as the density of the ϕ_j values is only half of that of the rigorous solution. But as the solution (4′) is complex, each value j corresponds to two real solutions, representing two independent waves of equal velocity in opposite directions. Each frequency is therefore to be counted twice, and the total number is the same as for the correct solution.

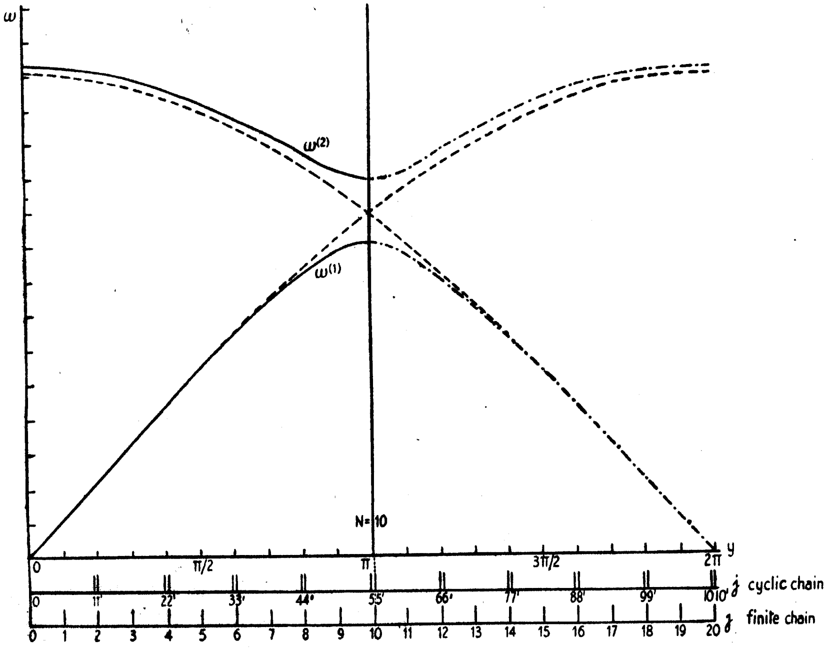


Figure 2.

The main point is the fact that the functions $\omega^{(1)}(\phi)$, $\omega^{(2)}(\phi)$ are exactly identical for the rigorous and the cyclic solutions. There are two branches, as illustrated in figure 2. This shows the well-known graph of the functions $\omega^{(1)}(\phi)=\sqrt{(2\alpha\epsilon^{(1)}/m)}$, $\omega^{(2)}(\phi)=\sqrt{(2\alpha\epsilon^{(2)}/m)}$ given by (9), and the ϕ_j-values, with $N=10$, for the rigorous and the cyclic solution.

The dotted lines represent the case of equal masses, $\mu=0$, where the frequencies are

$$\omega^{(1)}=2\sqrt{\frac{a}{m}}\,|\sin\phi/4|,\qquad \omega^{(2)}=2\sqrt{\frac{a}{m}}\,|\cos\phi/4|\,; \qquad \ldots\ldots(14)$$

these curves intersect at $\phi=\pi$, the end point of the domain of ϕ, which contains the whole spectrum just once. For different masses, $\mu\neq0$, the two branches deviate from these curves with increasing ϕ and differ for $\phi=\pi$ by a finite distance. The upper branch $\omega^{(2)}$ corresponds in 3-dimensional lattices to the "optical" vibrations, to which Raman attributes the new x-ray reflections and which he supposes to be monochromatic. In fact, they are spread over finite frequency intervals. In the linear model the upper branch becomes monochromatic if $m_2 \ll m_1$; then an expansion of (9) with respect to m_2/m_1 leads to the approximate values

$$\omega^{(1)}=\sqrt{\frac{\alpha}{m_1}}\,|\sin\phi/2|,\qquad \omega^{(2)}=\sqrt{2\frac{\alpha}{m_2}}, \qquad \ldots\ldots(15)$$

the second of which is constant. But even in this case Raman's assumption concerning the x-ray scattering cannot be justified, as I shall explain briefly in the last section.

The preceding considerations are an example of classical mechanics. As Raman very rightly insists that the diffuse x-ray spots are due to a quantum effect, it is desirable to formulate the results in such a way that they conform with orthodox quantum mechanics.

For this purpose the conjugate momenta p_l, p_l' of the displacements u_l, v_l must be introduced; the variables commute, except the pairs u_l, p_l and v_l, p_l', and we have, for instance,

$$u_lp_l-p_lu_l=\hbar i. \qquad \ldots\ldots(16)$$

Now we consider the functions (4) where A_1, A_2, B_1 can be expressed as multiples of B_2 (as explicitly given for B_1 by (12)), and determine B_2 by the normalization condition

$$\sum_l (m_1u_l^2+m_2v_l^2)=1. \qquad \ldots\ldots(17)$$

There are two such pairs of functions u_l, v_l for each integer j, corresponding to the two frequencies $\omega_j^{(1)}, \omega_j^{(2)}$; we write these

$$\begin{pmatrix} u_{l_j}^{(1)} & v_{l_j}^{(1)} \\ u_{l_j}^{(2)} & v_{l_j}^{(2)} \end{pmatrix}. \qquad \ldots\ldots(18)$$

If l and j assume the values 1, 2, ... N, these small matrices form a big matrix of order $2N$ which is orthogonal in the sense of the normalization (17).

We now introduce *normal coordinates* $\xi_j^{(1)}, \xi_j^{(2)}$, by the linear (orthogonal) transformation:

$$\begin{aligned} u_l&=\sum_j (u_{l_j}^{(1)}\xi_j^{(1)}+u_{l_j}^{(2)}\xi_j^{(2)}),\\ v_l&=\sum_j (v_{l_j}^{(1)}\xi_j^{(1)}+v_{l_j}^{(2)}\xi_j^{(2)}), \end{aligned} \qquad \ldots\ldots(19)$$

and the conjugate momenta $\eta_j^{(1)}, \eta_j^{(2)}$ by the same transformation of p_l, p_l'. Then the $\xi_j^{(k)}, \eta_j^{(k)}, (k=1, 2)$, satisfy the commutation laws

$$\xi_j^{(k)}\eta_{j'}^{(k')}-\eta_{j'}^{(k')}\xi_j^{(k)}=\delta_{jj'}\,\delta_{kk'}. \qquad \ldots\ldots(20)$$

The kinetic and potential energy

$$T=\tfrac{1}{2}\sum_{l=1}^{N}\left(\frac{1}{m_1}p_l^2+\frac{1}{m_2}p_l'^2\right),$$
$$U=\frac{\alpha}{2}\sum_{l=1}^{N}(v_l-u_l)^2+\frac{\alpha}{2}\sum_{l=1}^{N-1}(u_{l+1}-v_l)^2 \qquad \ldots\ldots(21)$$

are transformed into sums of squares

$$T=\tfrac{1}{2}\sum_{j=1}^{N}(\eta_j^{(1)2}+\eta_j^{(2)2}),$$
$$U=\tfrac{1}{2}\sum_{j=1}^{N}(\omega_j^{(1)2}\,\xi_j^{(1)2}+\omega_j^{(2)2}\,\xi_j^{(2)2}), \qquad \ldots\ldots(22)$$

where $\omega_j^{(1)}$, $\omega_j^{(2)}$, are the two branches of the frequency spectrum, defined by (9).

The dynamics of the crystal are described in quantum mechanics with the help of the Hamiltonian

$$T+U=H(\xi,\,\eta)=\sum_{j=1}^{N}\{H_j^{(1)}(\xi_j^{(1)},\,\eta_j^{(1)})+H_j^{(2)}(\xi_j^{(2)},\,\eta_j^{(2)})\}, \qquad \ldots\ldots(23)$$

where

$$H_j^{(k)}(\xi,\,\eta)=\tfrac{1}{2}(\eta^2+\omega_j^{(k)2}\xi^2), \qquad (k=1,\,2) \qquad \ldots\ldots(24)$$

is the Hamiltonian of a harmonic oscillator with the frequency $\omega_j^{(k)}$.

Analogous formulae will hold for any 3-dimensional lattice where l and ϕ represent triads of numbers and k may have more than two values.

§ 3. X-RAY SCATTERING

This section contains a short critical comparison of Raman's "theory of quantum reflections" with the correct quantum theory of scattering.

Raman starts with an expression for the amplitude of the scattered wave, of the form

$$E_0\,e^{i\omega t}\sum_{l}\alpha_l\,e^{i\Delta_l^0}\,.\,e^{i\Delta_l'}; \qquad \ldots\ldots(25)$$

here E_0, ω are amplitude and frequency of the incident wave, α_l the polarizability of the atom l, $\Delta_l^0=\frac{2\pi}{\lambda}r_l\,.\,(s'-s)$, λ the wave length, r_l the position vector of the atom l, s and s' the unit vectors in the direction of the wave normals of the incident and scattered beam. The last factor expresses the influence of the vibrations; Raman assumes that the Δ_l' are proportional to the displacements of the individual atoms and puts $\Delta_l'=u_l\cos(\omega_l t+\epsilon_l)$. This is not correct as the atoms are coupled; one has to put

$$\Delta_l'=\sum_{j}u_{l_j}\,\xi_j, \qquad \ldots\ldots(26)$$

where ξ_j are the normal coordinates, which perform harmonic vibrations. Correcting this point, we write

$$\xi_j=\cos(\omega_j t+\epsilon_j) \qquad \ldots\ldots(27)$$

and expand

$$e^{i\Delta_l'}=1+i\,\Delta_l'=1+i\sum_{j}u_{l_j}\cos(\omega_j t+\epsilon_j). \qquad \ldots\ldots(28)$$

If the cos is replaced by exponential functions of $\pm i(\omega_j t+\epsilon_j)$ and (28) substituted in (25) we see that, apart from a term with unchanged frequency, two terms with modified frequencies $\omega+\omega_j$ and $\omega-\omega_j$ and an arbitrary phase are obtained. These terms represent, according to Raman, the new quantum reflections,[†] and all his further conclusions are derived from these (*Symposium*, p. 338). But these formulae are quite classical.

The correct procedure according to quantum mechanics consists in considering the vibrating lattice in two vibrational states and forming the matrix element corresponding to the transition $n \rightarrow n'$; if the α_l are considered as independent of the vibrations (neglecting the Raman effect proper[‡]) only the Δ_l' depend on the ξ_j and hence on n, n'. Thus

$$E_0\, e^{i\omega t} \sum_l \alpha_l e^{i\Delta_l^0} \left(e^{i \sum\limits_j u_{lj} \xi_j} \right)_{nn'} e^{i(\omega_{nn'} t + \delta_{nn'})}, \qquad \ldots\ldots(29)$$

where $\hbar\omega_{nn'}$ is the energy difference between the states n, n' and $\delta_{nn'}$ an arbitrary phase. This shows that there are, apart from the unchanged (coherent) wave ($n=n'$, $\omega_{nn}=0$), a set of incoherent waves with changed frequencies $\omega+\omega_{nn'}$, one for each possible transition.

Actually this change of frequency is quite negligible in view of the enormous size of the x-ray photon $\hbar\omega$ as compared with infra-red quanta $\hbar\omega_{nn'}$. The important point is the incoherence, which implies that the intensity is obtained by adding the squares of the moduli of the individual wavelets. The total intensity is therefore obtained by multiplying the squared modulus of (29) by the Boltzmann factor corresponding to the initial state n, and summing first over all transitions from any other state n' to n, and then over all n. The first of these summations can be rigorously performed, since (with real eigenfunctions) for any matrix $f(n, n')$ one has

$$\sum_{n'} |f(n, n')|^2 = |f|^2(n, n').$$

In this way the intensity is found to be

$$I = I_0 \sum_{ll'} \alpha_l \alpha_{l'}^* e^{i(\Delta_l^0 - \Delta_{l'}^0)} \prod_j \text{mean}\left(e^{i\mu_j^{ll'} \xi_j}\right), \qquad \ldots\ldots(30)$$

where $\mu_j^{ll'} = u_{lj} - u_{l'j}$ and

$$\text{mean}\,(e^{i\mu\xi}) = \sum_n (e^{i\mu\xi})_{nn'} \frac{e^{-\epsilon_n/kT}}{\sum\limits_n e^{-\epsilon_n/kT}}. \qquad \ldots\ldots(31)$$

Here n means the quantum number of a single oscillator (amplitude ξ), and $\epsilon_n = \hbar\omega(n+\frac{1}{2})$ its energy.

This average of a harmonic function of the oscillator amplitude seems to me the central problem of the theory, as all consequent operations are of a purely geometric character. But the determination of this mean (31) is not quite trivial and has been the source of much confusion.

Debye (1913, 14) replaced $e^{i\mu\xi}$ by $1 + i\mu\xi - \frac{1}{2}\mu^2\xi^2 + \ldots$ and taking the mean term by term [mean $(\xi) = 0$, mean $(\xi^2) = \bar{\epsilon}/\omega^2$, where

$$\bar{\epsilon} = \hbar\omega\left(\frac{1}{2} + \frac{1}{e^{\hbar\omega/kT} - 1}\right) = \frac{\hbar\omega}{2} \coth\left(\frac{1}{2}\frac{\hbar\omega}{kT}\right) \qquad \ldots\ldots(32).$$

[†] As a matter of fact, Raman does not use the expansion (28) from the start, but a little more general expansion into a series of Bessel functions; but later he uses only the first terms of the power series of these functions, which is equivalent to the expansion (28).

[‡] This effect has been considered in one of my papers quoted, at least for the perfectly symmetrical modes which do not produce depolarization.

is Planck's oscillator energy] he obtained $1-\frac{1}{2}\mu^2\bar{\epsilon}/\omega^2+\dots$ and replaced this by an exponential function,

$$\text{mean}\,(e^{i\mu\xi})=e^{-\frac{1}{2}\mu^2\bar{\epsilon}/\omega^2}. \qquad \dots\dots(33)$$

Waller (1925) devoted an appendix of his important paper to an attempt at improving this procedure, but succeeded only for high temperatures (Boltzmann statistics). A rigorous proof was given by Ott in 1935, and a simpler one in the first of our papers (Born and Sarginson, 1941).

The correct derivation of (33) is not an unnecessary refinement,—for the exponent becomes infinite when $\omega\to 0$; hence Debye's expansion is valid only if very small values of ω can be excluded. But this is impossible, as the product appearing in (30) contains factors corresponding to every frequency, including those of the acoustical branches which approach zero; for these the single factors of the form (33) cannot be expanded.

But if the product is transformed,

$$\prod_j \text{mean}\,(e^{i\mu_j^{ll'}\xi_j})=\exp\Big(-\frac{1}{2}\sum_j \mu_j^{ll'}\bar{\epsilon}_j/\omega_j^2\Big)=\Pi_{ll'}; \qquad \dots\dots(34)$$

the sum in the exponent converges, even for the acoustical branches. The result can then, of course, be expanded into powers of the small quantities μ:

$$\Pi_{ll'}=1-\frac{1}{2}\sum_j (\mu_j^{ll'})^2\,\bar{\epsilon}_j/\omega_j^2+\dots. \qquad \dots\dots(35)$$

Substituting this in (30) we obtain two terms, one representing the Laue scattering, the other the influence of the thermal motion:

$$I=-I_0\cdot\frac{1}{2}\sum_j \frac{\bar{\epsilon}_j}{\omega_j^2}\sum_{ll'}\alpha_l\alpha_l^*(\mu_j^{ll'})^2\,e^{i(\Delta_l^0-\Delta_{l'}^0)}. \qquad \dots\dots(36)$$

The further development of this formula for a lattice is a straightforward process which will not be explained here. I have dwelt on the question of Debye's expansion in some detail as several modern authors (Zachariasen, 1940, 1; Jahn, 1941, 2) have failed to recognize the difficulty, and also because this appearance of ω^2 in the denominator is connected with Raman's erroneous conclusions.

Raman wishes to prove that the acoustical branch of the spectrum cannot produce sharp spots. Squaring his classical expression for the scattered amplitude, he identifies $\frac{1}{2}\omega^2u^2$ with kT; then he replaces ω by $2\pi c/\lambda$ (c = velocity of sound, λ = wave-length) and finds the intensity proportional to $kT\lambda^2$, a smooth function which cannot give rise to sharp maxima, but only to a diffuse background. But this is not correct. There is no constant velocity of elastic waves in the crystal, but a kind of dispersion. For a monatomic linear chain the formula is given in the introduction; it can be written

$$\omega=2\sqrt{\frac{\alpha}{m}}\left|\sin\phi/2\right|=2\sqrt{\frac{\alpha}{m}}\left|\sin\frac{\pi a}{\lambda}\right|=2\sqrt{\frac{\alpha}{m}}\left|\sin\frac{\pi q}{b}\right|;\quad b=\frac{1}{a}; \qquad \dots\dots(37)$$

here $q=1/\lambda$ is the wave number, $b=1/a$ the lattice constant of the reciprocal lattice, and

$$\phi=\frac{2\pi q}{\lambda}=\frac{2\pi q}{b}. \qquad \dots\dots(38)$$

The corresponding expressions for a diatomic chain (§1, equation (9)) are similar periodic functions of $\phi/2=\pi q/b$. It is true that all frequencies and motions of the particles can be obtained by restricting ϕ to the interval $0\leqslant\phi\leqslant\pi$ (hence $0\leqslant q\leqslant\frac{1}{2}b$, or $2a\leqslant\lambda\leqslant\infty$); but each of these waves is the longest sample in a whole set of "equivalent" waves, given by ϕ, $\phi\pm\pi$, $\phi\pm2\pi$, ... It is necessary to admit these shorter waves in order to express in a simple form the interference condition or, what is the same thing, the conservation law for momentum, according to de Broglie.

The discussion of the expression (36) for the intensity gives the result that enhancement of scattered wavelets from different lattice points by interference takes place if the wave vectors Q, Q' of the incident and scattered radiation and the wave vector q of an elastic wave in the crystal satisfy the condition

$$Q'-Q=q. \qquad \ldots\ldots(39)$$

The intensity can then be expressed in the form

$$I=S(Q'-Q), \qquad \ldots\ldots(40)$$

where $S(q)$, the scattering power, is a function in the reciprocal-lattice co-ordinates (q_1, q_2, q_3). To illustrate this, let us consider the reflections by a definite set of parallel equidistant lattice planes, and assume that these planes vibrate as a whole normal to themselves.

These vibrations will be the same as those of a chain of equal equidistant particles, and the frequency given by (37).

The interference condition (39) reduces to

$$Q\,(\sin\vartheta'-\sin\vartheta)=q, \quad \cos\vartheta'-\cos\vartheta=0, \quad \text{hence } \vartheta'=-\vartheta, \qquad \ldots\ldots(41)$$

where $Q=1/\Lambda$, the wave number of the x rays, $q=1/\lambda$, that of the elastic waves. It is clear that q cannot be restricted to the interval $0\leqslant q\leqslant b/2$. Taking Raman's approximate expression for the mean amplitude, the scattered intensity will be

$$I=S(2Q\sin\vartheta), \qquad \ldots\ldots(42)$$

where

$$S(q)=\frac{kT}{\dfrac{4\alpha}{m}\sin^2(\pi q/b)} \qquad \ldots\ldots(43)$$

is the scattering power, a function distributed over the whole reciprocal space q, which is condensed around the points of the reciprocal lattice,

$$q=0,\ b,\ 2b,\ 3b,\ \ldots$$

where it has sharp (infinite) peaks. These correspond to the Bragg reflections, as $2Q\sin\vartheta=nb$ is the same as $2a\sin\vartheta=n\Lambda$ (figure 3). If all vibrations in 3 dimensions are taken into account, the expression $S(q)=S(q_1, q_2, q_3)$ is more complicated, but the main features are the same: it is a function in the reciprocal space additively composed of contributions of the different branches of the elastic spectrum. The terms corresponding to the acoustical branches have sharp peaks (infinities) at the points of the reciprocal lattice (corresponding to the Laue spots), because the frequency approaches zero. If the scattering power were completely condensed at these peaks there would be only the Laue spots; but as it is spread around them in virtue of the temperature motion, there is an additional scattering, not confined to the selected directions (for given wave-length) of the Laue spots, but increasing in intensity on approaching these directions. The maxima are observed as extra spots. Their intensity depends very strongly on the wave-length and direction of the incident beam with respect to the lattice, since the domains of scattering power have characteristic anisotropic shapes different for different Laue spots. I shall return to this question later. Here we stress the point that the maxima are produced by the acoustical branches of the vibrational spectrum.

The high-frequency branches are always well above the zero level and will therefore contribute only a diffuse background with weak maxima midway between Laue points, traces of which seem to have been found by Lonsdale and Smith.[†]

The part played by the low- and high-frequency branches is just the opposite to that which Raman supposes.

For the high-frequency branches, Raman makes an attempt to introduce quantum theory in order to get the dependence on temperature. His reasoning is rather strange. He distinguishes two cases. If the number of vibrational quanta is diminished by the incident radiation, he assumes the energy given by the Planck formula without zero energy; but if this number is increased he adds to this $\hbar\omega$ (twice the zero energy). Then he adds these values and obtains an expression which is in fact $2\bar{\epsilon}$, where $\bar{\epsilon}$ is the Planck formula (32) with zero energy (*Symposium*, p. 350). The rigorous theory leads to the same temperature factor for all frequencies, including the low branches. The experiments of Venkateswaran (*Symposium*, p. 387 ff.), which are supposed to confirm this dependence on temperature, were made on extra spots well separated from the corresponding Laue spots; this means that the corresponding scattering frequency is rather far above the zero limit of the acoustical branch and of the same order of magnitude (infra-red) as the optical branches. All the experiments showed was that for carborundum the oscillations responsible for the extra spots were of higher frequency than those for a soft organic crystal, a result which the theory of lattice dynamics undoubtedly predicts, because of the difference of quasi-elastic force in the two substances. These measurements cannot be claimed as a confirmation of Raman's theory.

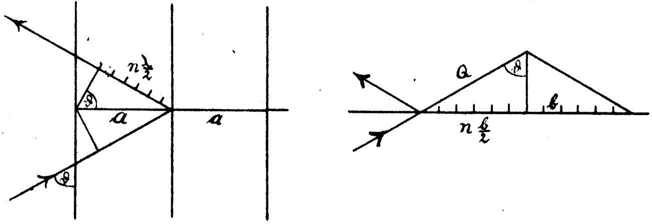


Figure 3.

The most condensed formulation of the result of the thermal theory is the following one[‡]:—The potential energy of the mutual forces between the lattice points is a quadratic form in terms of the displacements; for harmonic waves, the matrix of coefficients is a function of the wave numbers or of ϕ. In the case of a diatomic linear lattice (§ 2) we have, for instance,

$$D(\phi)=2\alpha\begin{pmatrix} 1 & -\frac{1}{2}(1+e^{-i\phi}) \\ -\frac{1}{2}(1+e^{i\phi}) & 1 \end{pmatrix}, \qquad \ldots\ldots(39)$$

[†] According to private information.

[‡] Materially contained in Waller's dissertation; the particular formulation with matrix functions is given in my last paper.

which is easily recognized as the matrix of the coefficients of the equations (5′).

The scattering power for a lattice composed of different atoms is also a quadratic form; the coefficients refer to the interaction of the pairs of different atoms and are also functions of ϕ. Let $S(\phi)$ be this scattering matrix. Then the result of the theory can be expressed in the form

$$S(\phi)=\frac{\hbar}{2}D^{-\frac{1}{2}}(\phi)\coth\left(\frac{\hbar}{2\boldsymbol{k}T}D^{\frac{1}{2}}(\phi)\right)=\boldsymbol{k}TD^{-1}(\phi)+\ldots \qquad \ldots\ldots(40)$$

The scattering matrix $S(\phi)$ is a function of the dynamical matrix $D(\phi)$, and just the same function that, according to (32), $\bar{\epsilon}(\omega)/\omega^2$ is of ω^2. Matrix equations like this are typical of quantum mechanics and could never be obtained by methods such as are used by Raman. The formula (40) may become of far-reaching importance, as it allows us to calculate $D(\phi)$ in terms of $S(\phi)$, and hence to determine the dynamical structure of the lattice from photographs of the extra spots. At present no quantitative data are available for this purpose. But the theory has been confirmed by considering the neighbourhood of the Laue spots, where the first term of the expansion, given by (40), is predominant and can be expressed with the help of the elastic constants. These formulae have been independently developed by Zachariasen (1940, 1), Sarginson (1942) and Jahn (1941, 2). The scattering power can be considered as condensed in regions surrounding the points of the reciprocal lattice, and Jahn has calculated the *iso-diffusion* surfaces. He found that for cubic crystals with marked elastic anisotropy, these surfaces are very far from being spherical, but have tentacles stretching out in different directions, and different for each lattice point. Knowing these surfaces, it is easy to find the scattered intensity with the help of Ewald's sphere of reflection (Ewald, 1913). This leads to characteristic differences for the intensities of the extra spots in the vicinity of different Laue spots, strongly dependent on the direction and wave-length of the incident beam. These predictions of the theory have been strikingly confirmed by experiments of Lonsdale and Smith (1942). Zachariasen also found good agreement of his formulae with observations (1940, 1).

Raman tries to explain the anisotropy of scattering power, deduced from the observations of his collaborators, by assuming that his monochromatic waves are propagated in selected directions. Rama Pisharoty (*Symposium*, pp. 377, 435) has carried out such calculations, but the assumptions made are rather arbitrary and the results not in good agreement with the facts (Lonsdale and Smith, 1942).

Finally I wish to mention another argument brought forward by Raman in favour of his assumption of monochromatic infra-red vibrations. His collaborator, Bisheshwar Dayal (*Symposium*, p. 421), has carried out calculations of the specific heat of metals in order to show that these cannot be represented by Debye curves, but that Einstein terms have to be added. This is quite true; the suggestion, however, is not new but already contained in my book *Dynamik der Kristallgitter* (1915) (formula 200, p. 77). As the optical branches are rather narrow, I showed that for quantities like the specific heat or the thermal expansion coefficient (which are average values over the whole spectrum) they may be replaced by monochromatic vibrations. This fact in no way supports Raman's hypothesis.†

§4. CONCLUSION

I wish to add a personal remark. It was Raman's first publication on the new scattering effect which directed my attention to it. His striking observations were obviously directed by his theoretical ideas. A theory which leads to a series of good experiments has fulfilled its first task. But its permanent value depends on the question whether it can be built into the system of well-established facts, and whether its predictions are confirmed by independent experiments. I tried to reconcile Raman's ideas with the results of quantum mechanics and lattice theory; but he, unfortunately, insists on his formulations. The present paper is intended to show why this is impossible. Scattering of x rays by atoms is well inside the boundary of established knowledge, not on the fighting front, like nuclei, mesons, etc., and it is extremely unlikely that anything quite unexpected should happen. Therefore Raman's expectations go too far. On the other hand, the story of the development of this subject is not very glorious for theoretical physics. It is rather similar to that of the optical Raman effect. Here, as well as there, the complete theory was given long before the phenomenon was observed; but in both cases the discussion of the formulae was so imperfect that no real predictions were made and no hints given to the experimenter. Yet after the discovery of the optical Raman effect the connection between observation and theory was established at once and acknowledged by Raman himself. May these pages help to produce the same effect for the x-ray scattering.

REFERENCES

Born, M., 1923. *Atomtheorie des festen Zustandes* (Teubner, Leipzig), p. 588.

Born, M. and von Kármán, Th., 1912. *Phys. Z.* **13**, 297.

Born, M. and Goeppert-Mayer, Maria, 1933. *Handb. d. Phys.* (Springer, Berlin), **24**, 11, chap. 4, p. 623; see in particular p. 638.

Born, M. and Sarginson, K., 1941. *Proc. Roy. Soc.* A, **179**, 69.

Courant, R., 1920. *Math. Z.* **7**, 1.

Courant, R. and Hilbert, D., 1931. *Methoden der mathematischen Physik*, **1** (Springer, Berlin), chap. vi, § 4.

Debye, P., 1913. *Verh. dtsch. phys. Ges.* **15**, 678, 738, 857; 1914. *Ann. Phys., Lpz.*, **43**, 49.

Ewald, P. P., 1913. *Phys. Z.* **14**, 465.

Faxén, H., 1923. *Z. Phys.* **17**, 266.

Jahn, H. A., 1941. *Nature, Lond.*, **147**, 511; 1942. *Proc. Roy. Soc.* A, **179**, 320.

Lonsdale, K. and Smith, H., 1941. *Proc. Roy. Soc.* A, **179**, 8.

Lonsdale, K. and Smith, H., 1942. *Nature, Lond.*, **149**, 21.

Ott, H., 1935. *Ann. Phys., Lpz.*, **23**, 169.

Raman, Sir C. V. and Collaborators, 1941. "Symposium of papers on the quantum theory of x-ray reflexions." *Proc. Ind. Acad. Sci., Bangalore,* A, **14**, 317 (quoted as *Symposium*), and *Proc. Roy. Soc.* A, **179**, 289, 302.

Sarginson, K., 1942. *Proc. Roy. Soc.* A, in the press.

Waller, T., 1923. *Z. Phys.* **17**, 348; 1925. *Dissertation* (Uppsala); 1928. *Z. Phys.* **51**, 213.

Weyl, H., 1911. *Math. Ann.* **71**, 441; 1912. *Ann. Phys., Lpz.*, **141**, 163; 1914. *Ibid.* **143**, 177; 1915. *Rendiconti Palermo*, **39**, 1.

Zachariasen, W. H., 1940. *Phys. Rev.* **57**, 597, 795; 1941. *Ibid.* **59**, 860, 909.

---

† When this MS. was just finished there appeared in the *Proc. Ind. Acad. Sci.* a new symposium of papers on the thermal energy of crystalline solids, in which this hypothesis is expanded and supported by experimental investigations on specific heat. My critical remarks are in no way affected by these new papers.

NOTE ON RAMAN'S THEORY OF THE SPECIFIC HEAT OF SOLIDS

By M. BLACKMAN, M.Sc., Ph.D.,
Imperial College, London

Received 27 *March* 1942

ABSTRACT. Raman's version of the theory of the specific heat of solids is examined; it is shown that the existing (lattice) theory of specific heat gives better agreement with experiment.

In a recent publication, Raman (1941) has proposed a reformulation of the theory of the specific heat of solids, rejecting both the Debye theory and the lattice theory of specific heats, (cf. Born, 1923). Raman's theory replaces the (in the main) continuous vibrational spectrum of the current theory by a few discrete lines—the specific heat being therefore calculated from a sum of a few Einstein specific-heat functions.

The arguments which led Raman to the rejection of lattice theory are of a purely qualitative nature; these have been dealt with by Born (1942).* This note is concerned with the results obtained, and to be expected, from Raman's theory.

In the majority of cases in which the theory is applied by Raman's collaborators (Norris, Dayal, Anand, Venkateswaran, 1941), the frequencies of vibration of the crystal considered are obtained "from a perusal of the specific-heat data"; the agreement obtained with experiment does not, therefore, prove more than that a suitable sum of Einstein functions can give a good fit to the specific-heat data for a solid. A very good demonstration of this fact was given, in 1911, by the success of the Nernst-Lindemann formula.

In the particular case of rocksalt (Venkateswaran, 1941) the frequencies, with one exception, were obtained from the more or less clearly defined peaks in the Raman spectrum, on the assumption that these peaks are octaves of the required lattice vibrations. *The remaining frequency is assumed.* The frequencies for sylvine (Venkateswaran, 1941) were all taken to be 0·78 of those of rocksalt. The resulting specific heats "agree reasonably well with experimental results", and it is concluded "the investigation thus clearly demonstrates that the present approach to the theory of specific heat should replace the earlier ones due to Debye, Born and others".

No account is given in the investigation of the results of lattice theory; in view of the importance of any new theory of specific heat, it seems desirable that such a comparison should be made. This is done below (table 1).

The experimental results are due to Clusius and Perlick (1934) (for rocksalt), and to Keesom and Clark (1935) (for sylvine). Specific heats on the basis of lattice theory have been calculated by Kellermann (1941) and Iona (1941) for rocksalt and sylvine, respectively. It should be noted that there are no adjustable constants in the lattice theory.

---

* I am indebted to Professor Born for the information that a paper on this point is in progress of publication in the *Proceedings of the Physical Society*. (See this Part, p. 362.—Ed.)

Table 1. Specific heats of rocksalt and sylvine

T° (K.)	C_v (expt.)	C_v (lat. theory)	C_v (Raman's theory)
Rocksalt			
10	0·0326	0·0328	0·036
20	0·310	0·317	0·28
30	1·150	1·154	0·96
40	2·380	2·380	2·04
Sylvine			
2·99	0·00206	0·0021	0·0016
4·01	0·0047	0·0049	0·0046
6·52	0·0191	0·0227	0·0219
8·00	0·046	0·043	0·043
10·06	0·084	0·087	0·0907

Judging solely by the agreement between theory and experiment, it is difficult to see any superiority on the part of Raman's theory.

There is a further point on which Raman's theory can be tested. The lattice theory of solids (Born, 1923) shows that the specific heat at very low temperatures can be calculated from the elastic constants. On Raman's theory, the specific heat at low temperature will depend on certain low-frequency lattice vibrations and *not* on the magnitudes of the elastic constants. This is, therefore, a point on which the theories disagree, and one which can be tested by reference to experiment.

Raman does consider the question briefly in his discussion of the Debye theory (Raman, 1941, p. 463), but dismisses the agreement as unimportant—" as was pointed out in Einstein's earliest papers, however, the existence of approximate relations between the atomic frequencies and elastic constants is indicated by elementary considerations ".

One would not expect " approximate relations " to lead to reasonably good agreement in all cases in which a real test is possible. There are only a few cases in which the specific heat has been measured to sufficiently low temperatures and for which, in addition, elastic data are available. The most important of these are sylvine and rocksalt, the thermal data for which have been quoted above; the elastic data have been given by Durand (1936).

The comparison is made in terms of the constant θ calculated independently from the thermal and the elastic data; the temperature at which the comparison is made is 10° K. for NaCl (the lowest temperature available) and 2·99° K. in the case of KCl.

Table 2

	θ (sp. heat)	θ (elastic)
NaCl	230	246
KCl	308	320

The agreement is much closer than would be expected from an " approximate relationship ", and there are reasons for believing that the agreement is better than is indicated in table 2 (cf. Blackman, 1942).

A further case which can be cited is that of zinc, for which a value of θ (elastic) of 320 is obtained (cf. Eucken, 1929). The value of θ (sp. heat) is as low as 195 at 20° K., but rises steeply at lower temperatures (Keesom and v. d. Ende, 1932), and if the electronic specific heat is allowed for, eventually reaches a value of about 310 at 4° K. This is a particularly interesting case because of the large difference between the value of θ (sp. heat) at 20° K. and at 4° K.

The evidence presented above points to a very close correspondence between lattice theory and experiment. In view of this, quite apart from other reasons, the claim that Raman's theory " should replace the earlier ones by Debye, Born and others " seems, to put it at its highest, somewhat premature.

REFERENCES

Anand, 1941. *Proc. Ind. Acad. Sci., Bangalore*, A, **14**, 483.
Blackman, 1942. *Rep. Progr. Phys.* **8**, 20.
Born, 1923. *Atomtheorie des festen Zustandes* (Leipzig).
Clusius and Perlick, 1934. *Phys. Z.* **35**, 943.
Dayal, 1941. *Proc. Ind. Acad. Sci., Bangalore*, A, **14**, 473.
Durand, 1936. *Phys. Rev.* **50**, 453.
Eucken, 1929. *Handb. d. Exp. Phys.* **8**, 1. 243.
Iona, 1941. *Phys. Rev.* **60**, 824.
Keesom and Clark, 1935. *Comm. Leiden*, no. 232 d.
Keesom and v. d. Ende, 1932. *Comm. Leiden*, no. 19 b.
Kellermann, 1941. *Proc. Roy. Soc.* A, **178**, 17.
Norris, 1941. *Proc. Ind. Acad. Sci., Bangalore*, A, **14**, 499.
Raman, 1941. *Proc. Ind. Acad. Sci., Bangalore*, A, **14**, 459.
Venkateswaran, 1941. *Proc. Ind. Acad. Sci., Bangalore*, A, **14**, 506.

NOTE ON THE COMPARISON OF THE SUBSIDIARY MAXIMA IN ELECTRON DIFFRACTION AND X-RAY DIFFRACTION PATTERNS FROM SINGLE CRYSTALS*

By A. CHARLESBY, Ph.D., D.I.C.,
Imperial College of Science and Technology, S.W.7

MS. received 27 *March* 1942.

ABSTRACT. The occurrence of diffuse bands in electron diffraction and their apparent non-appearance in x-ray diffraction is accounted for by differences in wavelength, intensity, crystal size, etc. Under suitable conditions it is expected that they may be observed with x rays.

A qualitative comparison is made of the Faxén-Waller theory, used to explain extra spots in the x-ray diffraction pattern, with the hypothesis advanced to account for the presence of diffuse bands. The former is in many respects equivalent to assuming

---

* This note arose out of the discussion on the lecture by Mrs. Lonsdale, to whom the author is grateful for helpful discussions.

approximate coherence in thermal vibrations over very limited regions of the crystal. If these regions shrink to unit-cell dimensions, the molecules therein vibrate very nearly independently, giving rise to diffuse bands, in addition to the usual Laue or Bragg spot patterns.

---

§ 1. INTRODUCTION

In the study of single crystals of some organic compounds by electron diffraction, we (Charlesby, Finch and Wilman, 1939) have observed very diffuse maxima, in addition to the normal Laue spot pattern. These maxima are related to the orientation of the molecules of the crystal relative to the incident beam, and their appearance can best be explained in terms of thermal motion. Assuming that groups of atoms (molecules) vibrate independently of their neighbours, the expression derived for the diffracted intensity was found to be

$$I \propto (A^2+B^2)\frac{\sin^2 M_1 h\pi}{\sin^2 h\pi}\cdot\frac{\sin^2 M_2 k\pi}{\sin^2 k\pi}\cdot\frac{\sin^2 M_3 l\pi}{\sin^2 l\pi}(\overline{\cos\phi})^2 + N\sum_p (A_p{}^2+B_p{}^2)(1-\overline{\cos\phi^2}).$$

The first term is the usual expression for the Laue spots modified by the structure factor and by a temperature factor $(\overline{\cos\phi})^2$, the value of which decreases as the angle of diffraction increases. The other term corresponds to the diffuse maxima, the intensities of which are modified by the structure factor of each independent group and by a temperature factor, which increases with the distance from the central spot. If, moreover, we assume that the restoring force on any molecule is proportional to the displacement, a more precise expression can be obtained for $(\overline{\cos\phi})^2$, viz.

$$(\overline{\cos\phi})^2 = \exp.\ \{[-16\pi^2 \sin^2\theta/\lambda^2](\boldsymbol{k}T/f)\},$$

where f is the force constant, 2θ the angle of diffraction and the other terms have their usual significance. These diffraction maxima are of considerable importance in the elucidation of the structure of organic compounds, and it is surprising that no similar phenomenon has so far been observed in X-ray work.

This failure can be attributed to several reasons:—

1. *Difference in wave-length*

In electron diffraction the wave-length used is of the order of 0·05 A. In Robertson's work (1933) on anthracene the Cu Kα wave-length of 1·54 A. was used. The diffraction maxima occur approximately at angles 2θ defined by the equation

$$2d\sin\theta = n\lambda.$$

In molecules of the type of anthracene, the maximum effective grating spacing is 1·22 A., although less intense maxima occur at spacings corresponding to 2·31 A. The first-order diffraction maxima would then appear at an angle 2θ such that

$$\sin\theta = \frac{\lambda}{2d} = \frac{1{\cdot}54}{2\times 1{\cdot}22} = 0{\cdot}63.$$

This implies that the angles of incidence and reflection from the corresponding grating spacings are equal. If this is not the case, the diffuse maxima occur at greater angles.

Robertson's experiments extend to angles 2θ such that $\sin\theta=0{\cdot}767$ (corresponding to the 019 diffractions), but this is an unfavourable direction for the observation of diffuse maxima. More favourable conditions prevail when the crystal is oriented to obtain the $40\bar{9}$, $40\,\overline{10}$, $60\bar{9}$ and $60\,\overline{10}$ diffractions, when the orientation of the molecule is such as to present the maximum effective grating spacing. The corresponding values of $\sin\theta$ observed by Robertson were 0·620, 0·690, 0·658, 0·709 respectively, and these appear to be the highest values of $\sin\theta$ examined with this position of the molecule relative to the beam. Thus, even in the most favourable case, the first-order diffuse bands would be barely visible at the edge of the photographic plate, possibly further decreased in intensity by the oblique incidence. Higher orders, which are relatively more intense, would not appear. The less intense subsidiary maxima, corresponding to a minimum effective spacing of 2·31 A., should, however, be observable. As they are of much lower intensity and, moreover, are greatly weakened by the thermal factor, they may readily be overlooked.

In our electron-diffraction patterns, the maximum value of $\sin\theta$ observable on the plate was greater than 1/15, so that the third order could be recorded if the intensity were sufficient.

The absence of the zero-order reflection still remains to be explained. This order would be evidenced by a diffuse band surrounding the undeviated beam. According to the theory mentioned above, the temperature factor is such that this maximum should be well-nigh eliminated. It seems probable that any other explanation for the diffuse maxima would lead to a somewhat similar result. Unfortunately, the electron-diffraction pattern near the undeviated beam is blackened by inelastically scattered electrons, and observation is difficult; nevertheless it does appear that some such decrease takes place in its neighbourhood. The corresponding increase in intensity with angle of deviation, deduced from the formula, can be readily observed.

2. *Difference in relative intensities*

Compare the relative intensities of diffuse maxima and spot patterns obtained by the x-ray and electron-diffraction methods. In the former, the exposures are generally arranged to give suitable blackening of the Bragg spots, on which the analysis is based. In the latter case, most of the spots observed are Laue spots and only two of the three Laue conditions are accurately fulfilled. The very intense electron beam and its greater scattering ensure that a few seconds exposure leads to suitable blackening of these relatively weaker spots, and in most cases to considerable over-exposure of the few Bragg reflections. In electron diffraction it is found that a correct exposure for the Laue-spot pattern gives satisfactory intensities for the diffuse bands. In x-ray work one would therefore require a considerable over-exposure of the Bragg spots before the diffuse bands became visible. It is only recently that long exposures have been used to study the x-ray type of extra spots; thus the failure to observe diffuse bands is understandable.

This argument assumes, of course, that the relative intensities of the diffuse maxima and Bragg spots are independent of crystal size. We formerly considered (Charlesby, Finch and Wilman, 1939) that the maximum intensity in a diffuse band is proportional to N, the number of unit cells in the crystal, whereas the intensity maximum of a Bragg spot is proportional to N^2; thus the diffuse maxima would be relatively more prominent in small crystals quite apart from the considerations outlined above. This statement appears to hold for our electron-diffraction patterns, but requires further qualification.

The intensity distribution in a spot depends not only on the Laue factors of the form $(\sin^2 M_1\pi h)/\sin^2 \pi h$, but also on other factors such as beam divergence and width. The Laue zones due to atom rows which are not too nearly parallel to the incident beam are generally so sharp that the spot size is essentially determined by the other factors involved. An increase in M_1, the number of unit cells in this direction (assumed already large), will only modify the form of the intensity distribution in the spot to a very slight extent. The intensity maximum of any given spot is in effect proportional to M_1.

When considering thick crystals a similar state of affairs holds for all directions in the crystal, so that the intensity is proportional to the product $M_1M_2M_3$, i.e. N. On the other hand, with thin crystal sheets, nearly perpendicular to the beam, M_1 is small for atom rows parallel to the beam; and for the range of diffracted beams recorded, the corresponding index h is always relatively small. A wide maximum occurs in the Laue factor and the reciprocal-lattice point is elongated in the direction normal to the sheet. This is evidenced in electron diffraction by a very wide Laue zone. As M_1 is increased, the intensity of spots near the maximum of the Laue zone increases as $M_1{}^2$, and at the same time the width of the Laue zone shrinks. In these thin specimens, therefore, the ratio of maximum spot intensity/maximum diffuse band intensity is proportional to thickness. The transition between the two forms of variation occurs for thicknesses such that the spot radius determined from the Laue conditions (i.e. $\sim 1/M_K$ of the distance between reciprocal lattice spots in the K direction) equals that inherently due to other causes. In electron diffraction this thickness may be estimated to be about 10^4 A.

Summarizing, the intensity of the diffuse bands should be proportional to the number of unit cells, N, whereas the intensity of a spot near the centre of a Laue zone (electron diffraction) varies as $M_1{}^2M_2M_3$ or NM_1 (where M_1 is the number of unit cells in a row nearly parallel to the incident beam). In X-ray work the observed Bragg-spot intensity varies as $M_1M_2M_3(=N)$. The ratio of the intensities of the diffuse bands to observed spots is then inversely proportional to thickness in electron diffraction, and is independent of thickness in X rays. Beam divergence in X rays will not affect this relationship. Similar conclusion can also be reached if one assumes the crystal under examination to consist of mosaic blocks, the dimensions of which are much smaller than, but independent of, the size of the crystal as used in X rays. In electron diffraction, in view of the thickness of the specimens used, the mosaic structure will occur in the plane of the film, but not at right angles to it.

3. *Other effects*

The relative increase in intensity of diffuse maxima in regions remote from the central spot can be readily seen in electron-diffraction patterns. If the bands due to only one anthracene molecule in the unit cell are considered, the number of principal maxima visible throughout the plate is eighteen up to the second order and thirty-six up to the third order. The zero order is eliminated by the thermal factor. There are, moreover, twice as many secondary maxima, whose intensity is 54/196 that of the primary maxima. These maxima placed in regular array are far more readily observed than the few which might appear in X-ray patterns (using a wave-length of 1·54 A.) near the edge of a steeply inclined plate. This difference will be accentuated by the relatively large angular spread of each diffuse band, caused by the very limited extent of the molecular grating. This difference in scale would then give rise to bands, each extending over one quarter of the pattern. It is doubtful whether these bands could be observed above the general background of the Weissenberg photographs used by Robertson.

It has also been suggested that the diffuse bands may be due to some surface effect. It appears to me doubtful whether in this case the orientation which occurs in the bulk of the crystal would be preserved, as is shown by the patterns.

X-ray investigations on benzil have been carried out in an attempt to obtain diffuse maxima. It is doubtful whether positive results could be achieved, since, apart from the considerations outlined above, the structure of the molecule is not sufficiently regular. Any diffuse maxima observed would be correspondingly weak and broad. With anthracene it is possible to obtain diffracted beams in which all atoms of a molecule are in phase. The diffracted intensity is then fourteen times that due to the corresponding number of atoms considered independently. In benzil this ratio is probably not greater than 3 or 4, since it does not appear likely that the atoms of the two benzene groups will all be in phase simultaneously. An approximate calculation gives an estimated band breadth of about three times that to be expected from anthracene.

To attempt to investigate the diffuse-band phenomena with monochromatic X rays, one would require a very short wave-length, very long exposures, and a molecule of considerable symmetry (e.g. anthracene, pyrene) suitably oriented.

§ 2. COMPARISON OF ELECTRON DIFFUSE MAXIMA WITH X-RAY EXTRA SPOTS

Although experimental difficulties have so far prevented direct investigation of the effect of temperature on the diffuse band pattern, it appears that their appearance is, in fact, due to thermal vibrations. If this is the case, one would expect a close relation between the two types of subsidiary maxima present in single-crystal diffraction patterns, both of which should be predictable from a single comprehensive thermal theory. At the moment this theory has not been advanced, due both to the mathematical difficulties involved and to the lack of sufficiently accurate experimental data. We are therefore faced with the difficulty that, to account for these presumably related phenomena, we have to use two different explanations, the one based on vibrations throughout the crystal, the other on a hypothetical independence of phase in the vibrations of adjacent molecules. In what follows, an attempt is made to indicate possible lines along which the two apparently distinct explanations may be reconciled.

An interesting comparison may be drawn in terms of the phase relationship of the neighbouring diffracting elements or groups. Each scattering centre (e.g. atom) will be taken to have co-ordinates of the form $n_1+x_0+x_1$, $n_2+y_0+y_1$, $n_3+z_0+z_1$ along the three axes measured in terms of the corresponding axial distances a_1, a_2, a_3. n_1, n_2, n_3 are integers defining the unit cell in which the atom is placed, x_0, y_0, z_0 are the additional fractional co-ordinates taken relative to the origin of the corresponding unit cell, and x_1, y_1, z_1 are the additional displacements due to thermal vibrations.

If *h, k, l* are the Laue numbers for given incident and emergent beams, the amplitude of the wavelet scattered from the atom can be written proportional to

$$fe^{i\omega t}\, e^{2\pi i\Sigma h(n_1+x_0+x_1)},$$

where the summation extends over the three co-ordinates and Laue indices. The amplitude for the whole crystal is then

$$A \propto e^{i\omega t} \sum_{n_0\ldots} \{e^{2\pi i\Sigma hn_0} \sum_{x_0\ldots} [fe^{2\pi i\Sigma hn_1}e^{2\pi i\Sigma hx_1}]\}$$
$$= e^{i\omega t}\, \Sigma\, \{e^{2\pi i\Sigma hn_1}\Sigma[fe^{2\pi i\Sigma hx_0}(\cos\phi + i\sin\phi)]\},$$

where ϕ is the additional phase-angle introduced by thermal vibration. $\phi = 2\pi i\Sigma hx_1$ and its average value is zero.

To obtain the mean intensity we must multiply A by its conjugate complex and average the resultant intensity over all possible vibrations of the crystal. The same result is achieved by taking all the cross-product terms arising from A and A^*, averaging, and then adding. This inversion of the order of summation and averaging is of course only permissible since we are dealing with a sum.

A typical cross-product term is

$$e^{2\pi i\Sigma h(n_1-n_1')}ff'e^{2\pi i\Sigma h(x_0-x_0')}(\cos\phi\cos\phi' + \sin\phi\sin\phi').$$

It is the average value with respect to time, of the last factor involving ϕ, which decides the effect of the thermal vibrations on the contribution of this cross-product term to the intensity.

An accurate evaluation of $\cos\phi\cos\phi' + \sin\phi\sin\phi'$ or $\cos(\phi-\phi')$ requires considerable information on the interatomic forces in the structure. Certain interesting conclusions can, however, be reached from more general considerations.

We have previously assumed (Charlesby, Finch and Wilman, 1939) that atoms vibrate within the same molecule in phase, but those in different, albeit neighbouring, molecules, are independent. With this approximation $\cos(\phi-\phi')$ has an average value $\overline{\cos\phi}\;\overline{\cos\phi'} + 0$ or $(\overline{\cos\phi})^2$ for cross products involving atoms in different molecules. For atoms within the same molecule, the mean value has the greater value, 1. The expression for the (average) intensity is therefore

$$I \propto \sum_{x_0}\ldots\sum_{x_0}{}'\ldots e^{2\pi i\Sigma h(n_1-n_1')}ff'e^{2\pi i\Sigma h(x_0-x_0')}(\overline{\cos\phi})^2 + \sum_{{}_1=n_1'}\ldots\sum_{x_0} ff'e^{2\pi i\Sigma h({}_0-x_0')}\times 1,$$

where the first summation excludes cross products involving the same molecule twice, while the second only includes these. The former restriction can be removed:

$$I \propto \sum_{n_1}\ldots\sum_{x_0}\ldots e^{2\pi i\Sigma h(n_1-n_1')}\, ff'e^{2\pi i\Sigma h(x_0-x_0')}\,(\overline{\cos\phi})^2$$
$$+ \sum_{n_1=n_1'}\ldots\sum_{x_0x_0'}\ldots ff'e^{2\pi i\Sigma h(x_0-x_0')}(1-\overline{\cos\phi}^{\,2}).$$

The latter summation arises of course from the relatively higher contribution of cross products from the same molecule, which are always in phase.

It is implicitly assumed in this argument that all molecules are subject to similar forces, and hence have a similar mean value for $\cos\phi$.

The expression for the intensity can then be readily seen to consist of the usual expression for the spot pattern, but modified by the thermal factor $(\overline{\cos\phi})^2$, and in addition, an expression corresponding to a very limited lattice, extending over one molecule only. The latter is multiplied by the thermal factor $(1-\overline{\cos\phi^2})$ and the number of terms of this nature, i.e. N.

$$I \propto (A^2+B^2)\,\Pi\,\frac{\sin^2 M_1 nh}{\sin^2 \pi h}\,.\,(\overline{\cos\phi})^2 + N\sum_p (A_p{}^2+B_p{}^2)(1-\overline{\cos\phi^2}),$$

in the case where there is more than one independently vibrating molecule per unit cell.

In a further approximation, it becomes necessary to consider the interaction between the displacement of different vibrating units (e.g. molecules). This may be carried out rigorously by resolution of thermal displacements into normal vibrations throughout the crystal. In more complicated crystals, such as those which give rise to diffuse bands, not all the information required is available. It can, however, be expected that the influence of the displacement of one molecule on its neighbours will be more considerable than that on more distant molecules. Thus ϕ and ϕ' are independent for sufficiently distant molecules, but related in some as yet undefined manner for adjacent molecules.

These two conceptions are in a sense complementary. In the thermal-wave conception, the displacement of any molecule (or a corresponding grouping) is obtained by a summation of the wave amplitudes with arbitrary phase angles. Since these waves form an almost continuous spectrum of frequencies over a wide range, coherence in the phase of the vibrations will only occur between molecules which are sufficiently close, e.g. when the distance apart is of the order of the shortest wave-length present in the vibrational spectrum. Around each vibrating unit we may therefore consider a rather indefinite region outside which thermal vibrations bear no phase relationship to that of the unit considered. Cross-product terms involving it and any outside units will therefore be modified by a thermal factor $\cos(\phi-\phi')$, where ϕ and ϕ' are independent, and whose average value, as above, is therefore $\overline{\cos\phi\cos\phi'}$, i.e. $(\overline{\cos\phi})^2$. On the other hand, cross products involving it and units falling within the region will give rise to different thermal factors $\cos(\phi-\phi')$. In particular, for the cross product of any molecule with itself, this factor has the value unity. The actual evaluation of the intermediate cases of partial dependence would entail methods and knowledge similar to those required in the normal vibration method. It can, however, be assumed in almost all cases that ϕ and ϕ' tend to be of the same sign, i.e. $\overline{\phi\phi'}>0$. In this case we have

$$1 \geqslant \overline{\cos(\phi-\phi')} > (\overline{\cos\phi})^2,$$

or $\overline{\cos(\phi-\phi')}=(\overline{\cos\phi})^2+T^2$, where T^2 generally increases from zero, as the two molecules which give rise to the cross-product term considered are closer.

The expression for the intensity is then obtained by a summation of all the cross-product terms, of which those corresponding to closer molecules are in general of larger amplitude.

$$I \propto (\overline{\cos\phi})^2 \sum_{n_1\ldots}\ \sum_{x_0\ldots} e^{2\pi i \Sigma h(n_1-n_1')} ff' e^{2\pi i \Sigma h(x_0-x_0')} + \sum_{n_1\ldots} \sum_{x_0\ldots} e^{2\pi i \Sigma h(n_1-n_1')} ff' e^{2\pi i \Sigma h(x_0-x_0')} . T^2.$$

The first term gives rise to the usual Laue-spot pattern modified by the thermal factor $(\overline{\cos\phi})^2$, while the second term will be similar in character, but the summation is, in fact, limited to regions of mutual coherence. The diffraction pattern obtained will apparently consist of the superposition of the usual spot pattern and approximately of the pattern obtained from a limited lattice, although the intensity distribution as between the spots will be modified.

It has already been shown (W. H. Bragg, 1941; Preston, 1939) that the extra spots observed in x rays can to a considerable extent be correlated with the existence of hypothetical groups of particles. This is in approximate agreement with the conclusions derived above. The assumption of groups is almost equivalent to taking T^2 as constant over a small region, surrounding each scattering centre, and zero over the remainder of the crystal. Actually it will decrease from the value $1-\overline{\cos\phi}^2$ to zero in a more continuous fashion. This distinction becomes important when attempts are made to deduce the size of the hypothetical group. The intensity distribution within each diffuse spot will be appreciably different in the two cases.

Moreover, it offers a means of establishing the approximate size of the *regions of coherent vibration*. In the case of certain simple inorganic structures (KCl, Al, etc.) these appear to cover not more than several unit cells in any direction. In the case of Al, Preston (1939) finds that this region extends over the twelve neighbours of any atom.

Returning now to the more complex structure (e.g. organic crystals), it may be expected that these regions will be relatively smaller, due to the larger number of vibrations to be considered, the greater amplitudes of vibration (cf. elastic constants), and the larger mass of each vibrating unit (resulting in slower vibration and increased importance of the higher frequencies). The effective coherent-vibration region may well cover only the unit cell. This, too, is borne out by consideration of the anthracene crystal, where a theory assuming virtual independence in the phase of the vibrations of different molecules leads to very satisfactory conclusions. It should perhaps be pointed out that the normal co-ordinate method, as given in the Faxén-Waller theory, is liable to fail when the amplitudes of vibration become very considerable (as is indicated by the low melting point of these crystals, etc.). The general arguments used above and the conclusions derived from them will not, however, be affected.

§ 3. CONCLUSION

Subject to suitable crystals being obtainable, the molecules of which can behave as diffraction gratings of sufficient extent, it is expected that diffuse band phenomena should be observable in x-ray diffraction. Wide angles of diffraction and short wave-lengths will, however, be required, as well as long exposures and suitable crystal orientation. Moreover, it is assumed that other effects (e.g. general scattering) do not obliterate the effects which might otherwise be noted.

Extra spots can be considered to arise from the lack of coherence in the thermal vibration of different atoms or molecules of the crystal. The lack of phase relationship between the periodic displacement of distant scattering centres causes a decreased contribution to the scattered intensity. On the other hand, the contributions from neighbouring vibrating atoms or molecules, where this lack of coherence is less marked, will be relatively larger. The total diffracted intensity may thus be considered to consist of two parts, the first arising from the whole crystal, but with a decreased intensity due to thermal agitation. The second part arises from the additional contribution from close neighbours in the structure; this part simulates the pattern obtained from very small groups of molecules, the size of which is determined by the region of coherent vibration, i.e. the region where the relationship between the phases of the vibrations of different atoms is not entirely random. This region appears to extend approximately over several atomic spacings in simple inorganic structures, and over a single unit cell in more complex organic molecular compounds.

REFERENCES

BRAGG, Sir W. H., 1941. *Proc. Roy. Soc.* A, **179,** 51.
CHARLESBY, A., FINCH, G. I. and WILMAN, H., 1939. *Proc. Roy. Soc.* A, **51,** 479.
PRESTON, G. D., 1939. *Proc. Roy. Soc.* A, **172,** 116 ; 1941. *Ibid.* **179,** 1.
ROBERTSON, J. M., 1933. *Proc. Roy. Soc.* A, **140,** 79.

RECENT REPORTS AND CATALOGUES

Production Control, by A. R. JACKSON. Pp. 16. KARDEX DIVISION, Library Bureau, Ltd., 1 New Oxford Street, London, W.C. 1. 1*s.*

Electric Furnaces, Type T. (Leaflet 51.) Pp. 4. JOHNSON, MATTHEY & CO., Ltd., 73/83 Hatton Garden, London, E.C. 1.

Drum-Reading Plunger-Type Colorimeter. (January 1942.) Pp. 8. THE TINTOMETER, Ltd., The Colour Laboratory, Salisbury.

Optical Glass Cells. (Catalogue L. 95–L. 95A.) Pp. 12. THE TINTOMETER, Ltd., The Colour Laboratory, Salisbury. .

Testing of Glass Volumetric Apparatus, by E. L. PEFFER and GRACE C. MULLIGAN. (National Bureau of Standards Circular C 434, December 1941, superseding C 9.) Pp. 27. U.S. DEPARTMENT OF COMMERCE, Washington. 10 cents.

American Standard Specification for Dry Cells and Batteries. (National Bureau of Standards Circular C 435, February 1942, superseding C 414.) Pp. 15. U.S. DEPARTMENT OF COMMERCE, Washington. 10 cents.

Low-cost Glazes for Structural Clay Products, by R. T. STULL and P. V. JOHNSON. (National Bureau of Standards Circular C 436, March 1942.) Pp. 20. U.S. DEPARTMENT OF COMMERCE, Washington. 10 cents.

REVIEWS OF BOOKS

The Theory of Rate Processes, by SAMUEL GLASSTONE, KEITH J. LAIDLER and HENRY EYRING. Pp. viii+611. (New York and London: McGraw-Hill Book Co., Inc., 1941.) 42*s*.

This book has the more explanatory sub-title *The Kinetics of Chemical Reactions, Viscosity, Diffusion and Electro-chemical Phenomena*. It is an account of the method, generally associated with the names of Eyring, Polanyi and others, of the calculation of rates of reaction from the fundamental properties of the reacting species.

Some of the first reactions to be carefully investigated were homogeneous gas reactions which proceeded, in the range accessible to measurement, at rates which could be accounted for quite simply on the basis of a kinetic-theory collision mechanism, assuming that only collisions fulfilling certain energetic requirements led to reaction. Further work led to the discovery of reactions with rates both considerably slower and faster than was predicted by the simple calculation of fruitful collisions. The theory of absolute reaction rates (or the transition-state method), developed later, shifts the emphasis from the collision to the activation mechanism. It is perhaps important to notice here that at the roots of both theories lies the assumption that the equilibrium concentration of "activated complexes" is unaffected by the progress of the reaction. The authors state this explicitly (p. 185), but they do not mention that there are certain reactions of very low activation energy, e.g. the quenching of the fluorescence of dissolved hydrocarbons by oxygen, and possibly certain chain reactions, where the assumption almost certainly does not hold.

The rates of many reactions may be expressed in the familiar form, rate$=A\,.\,e^{-E/RT}$, where A is a frequency factor and E is an energy of activation. The application of quantum mechanics to the evaluation of the energy of activation was first considered by London in 1928. His suggestion that a single function should be capable of representing the state of an electron throughout chemical reaction led to the derivation of an approximate equation giving the potential energy of a system with variation in the internuclear distances. The exact solution of this equation, even for the simplest cases, is exceedingly difficult, but by making further approximations—mainly regarding the distribution of energy between coulombic and exchange terms—Eyring and Polyani developed what has become known as the "semi-empirical method" for calculating activation energies. Their argument makes use of the concept of the energy barrier: considering the simple reaction, $A+B=AB^*=$ products, if the potential energy of the system is plotted as a function of an internuclear distance co-ordinate, a humped curve is obtained in which the maximum corresponds to the potential energy of the activated complex AB^*, the activation energy being given by the height of this maximum over the point corresponding to the system $A+B$ at infinite separation. To determine the rate of the reaction, it is then only necessary to evaluate the frequency with which molecules pass over the barrier. It is an important result of the transition-state theory that it is able to identify this frequency with the product of a universal frequency, $\boldsymbol{kT/h}$, and an equilibrium constant, expressed in terms of partition functions, so that A, the frequency factor, is identified with $\left(\frac{\boldsymbol{kT}}{\boldsymbol{h}}\right)\frac{F_*}{F_A F_B}$, F_*, F_A and F_B being the partition functions of the activated complex and of the reactants A and B respectively. (In practice, a constant, κ, unity or less, the *transmission coefficient*, dependent on the detailed shape of the potential energy contours in the neighbourhood of the activated state, has also to be introduced into this expression.)

The relation giving the rate of reaction may now be written :

$$\text{rate} = \kappa \left(\frac{\boldsymbol{k}T}{\boldsymbol{h}}\right) \frac{F_*}{F_\mathrm{A}F_\mathrm{B}} e^{-E_0/RT},$$

whence it can be shown that

$$\text{rate} = \kappa \left(\frac{\boldsymbol{k}T}{\boldsymbol{h}}\right) e^{-\Delta F_* R/T} = \kappa \left(\frac{\boldsymbol{k}T}{\boldsymbol{h}}\right) e^{-\Delta H_*/RT} e^{\Delta S_*/R},$$

where ΔF_*, ΔH_* and ΔS_* are the free energy, heat and entropy changes accompanying the activation process : here ΔH_* differs only slightly from the experimental activation energy. It is by no means an inconsiderable virtue of this treatment that emphasis is thus laid on the free energy, rather than on the heat, of activation. The rather general form of these expressions for the rate of passage over an energy barrier accounts for the possibility of the interesting applications of the theory noted in the sub-title of the book.

The book divides into two parts. Pp. 1 to 201 contain a rather academic introduction, and in Chapters II to IV, entitled *Quantum Mechanics*, *Potential Energy Surfaces* and *Statistical Treatment of Reaction Rates*, respectively, are derived the results to be used in the later application of the theory. The first part, therefore, is an exposition of the method outlined above : perhaps the main criticism is that Chapter II, a very concentrated discussion of the results of quantum mechanics required for the subsequent treatment, appears to fall between two stools in being too terse for the more general reader, while redundant for the student familiar with the subject. This criticism may be outweighed by the advantage of having a development of the subject in the form required by the authors for their present purpose.

In the succeeding pages (pp. 202 to 599) are discussed the applications of the theory : the headings are *Homogeneous Gas Reactions*, *Reactions involving Excited Electronic States*, *Heterogeneous Processes*, *Reactions in Solution*, *Viscosity and Diffusion* and *Electrochemical Processes* (Chapters V to X). Most of the material has already appeared in the literature (although some has so far only been published in rather inaccessible theses), but the attempt at a connected story—even if the story is at present incomplete—is successful, and the result highly stimulating. The applications discussed in the first four of these chapters have provided the considerable experimental justification for the essential correctness of the theory, and its impetus to experimental work is made clear in these pages. An interesting example is the $H-CH_4$ complex, which calculation shows to possess a fair measure af stability.

The chapter on viscosity and diffusion is less satisfactory. The application of transition-state theory to the liquid state, in conjunction with the theory of holes in liquids, does not, at present, appear to be very profitable. The exuberance of the authors, however, carries them on to make statements which are, at the least, highly arbitrary. For example, regarding the variation of viscosity with pressure, curves are given (figure 124, p. 509) of this variation for pentane and for ethyl ether. The calculated curves are obviously diverging more and more from the experimental points with increase of pressure, so that at a pressure of 12×10^3 kg./cm^3 the figures for ethyl ether are 62 millipoise (observed) and 105 millipoise (calculated). It hardly seems possible, therefore, to " predict with fair accuracy the influence of pressure on the viscosity of liquids ".

In the last chapter, *Electrochemical Processes*, an interesting and again stimulating outline is given of the application of the theory to problems of ionic mobility and of electrode phenomena.

The printing and lay-out of the book are excellent, and it is remarkably free from serious misprints. A few minor errors have been introduced into the references, but they are not of sufficient importance to be noted here. In figure 108 (p. 429) the error in the legend to the original diagram of Livingstone has been perpetuated: reaction IV should be " $CH_3CO_2C_2H_5 + OH^-$ ". The subject index is not very complete, and it is possible that a separate index of reactions would make for greater ease in using the book. The system of referencing used, that of numbered footnotes, is rather inconvenient when the reader, as sometimes happens, is referred to a footnote number which may be a score of pages back. It would have been better if the references had been collected at the ends of each chapter. These are small things, but they would have increased the usefulness of a book certain to be widely read. For, in spite of lack of self-criticism and the rather biassed outlook of this book, its readers will be grateful to the authors for a highly interesting account of a theory which seems likely to make many further important contributions to the study of physico-chemical phenomena.

R. F. B.

Practical Acoustics and Planning against Noise, by HOPE BAGENAL. (London: Methuen and Co., 1942.) Price 7*s*. 6*d*.

The appearance of a book concerned in the main with post-war development causes elation in some, exasperation in others. Certain it is that in the topsy-turvy world which war brings in its train, scientists who have, directly or indirectly, served on committees for the abatement of nuisances like smoke and noise are now asked to exercise their wits to produce as much smoke or noise as possible. It is, accordingly, in the light of what we hope a post-war world may look like, that this little volume must be reviewed.

In effect, it is a smaller version of the book which Mr. Bagenal and Dr. Alexander Wood wrote some years ago, with most of the citations and calculations omitted. Added to this, that Mr. Bagenal has brought it up to date, notably in regard to later experience in broadcasting studio and theatre design, and we find a book more suitable in size and scope for the student-architect than the practitioner for whom the first was mainly intended. There is, in addition, more emphasis on the reduction of and insulation against noise in this than in the earlier book. Nearly half the book is concerned with what one might call (by analogy with the corresponding divisions in ballistic science) " external acoustics ", and rather more than half with " internal acoustics ". Actually it is the latter which is the prime concern of the architect, the former being the domain of the sociologist and the engineer.

The author, in the latter portion of the book, gives useful information on suitable designs for various types of buildings. His remarks on the difficult acoustic problems met in designing a church are particularly worth noting.

The physicist may occasionally be shocked by what he will feel are unorthodox uses of familiar technical terms, but since this book is meant for architects he has really no right to complain as long as the author defines the interpretation which he attaches to such words.

E. G. R.

Tables of Squares of all numbers less than 300. 1 page on card. (London: Scientific Computing Service, Ltd., 1942.) 1*s*.

THE PROCEEDINGS OF THE PHYSICAL SOCIETY

VOL. 54, PART 5 1 September 1942 No. 305

SOME PROBLEMS RELATING TO OPTICAL GLASS

By W. M. HAMPTON, Ph.D., B.Sc.

I.A.R.I. AGRI RES. NEW DELHI LIBRARY

Inaugural Lecture to the Optical Group. MS. received 30 *March* 1942

§ 1. INTRODUCTION

When I heard that it had been decided to form an Optical Group of the Physical Society I received the news with considerable pleasure, as I had always felt that the death of the old Optical Society was a great loss, both to science and the optical industry. I feel that I am speaking for the whole optical industry when I say that this new Group will fill a definite need, and I trust that all optical manufacturers will give it their utmost support.

After I had accepted your Secretary's invitation to deliver a lecture on optical glass I found myself in something of a quandary. This was not due to a lack of problems relating to optical glass which could be profitably discussed, but rather to the fact that, owing to the vital character of the optical-glass industry during a war, the restrictions imposed on me by considerations of national safety were such that practically all the points which could be of interest to you were prohibited. It is not too much to say that modern war cannot be fought without optical instruments, and clearly such instruments could not be made unless there was a sufficient supply of optical glass of suitable quality. To mention only a few cases, submarine periscopes, range finders, aerial cameras, all employ optical glass of the highest quality and have, in fact, called for the development of special types of glass to meet their particular needs. Very considerable developments have taken place in the manufacture both of the instrument and the glass, not only in this country but, as you know, in Canada as well. In view of the restrictions imposed by the considerations mentioned earlier, I decided to deal as fully as I was permitted with various specific problems, and for this purpose I considered four main difficulties which face any manufacturer of optical glass under war-time conditions. These are:—

1. The difficulties of supplying glass in sufficient quantity.
2. The difficulty of the supply of raw materials of the necessary quality.
3. The effect of heat treatment on the optical properties of the glass made.
4. The rather more general question of durability.

Under each of these four heads I have taken one particular point and discussed it, and that will form the basis of my paper.

§ 2. SUPPLY OF GLASS

A Government company, Research Enterprises Limited, was formed during 1940 to erect a factory in Toronto for the manufacture of various instruments, and in some cases the raw materials which were necessary for them. When it was decided, about October 1940, to commence the manufacture of optical glass there, Colonel W. E. Phillips, the President of the company, approached Chance Brothers & Co., Ltd., through the Canadian and British Governments, with the suggestion of co-operation. As a result of that approach I visited Toronto in November 1940 in order to ensure the most rapid development of the new undertaking. The result was that an arrangement was reached whereby the fullest information concerning the manufacturing technique was transferred to the Canadian factory, and certain physicists from Canada came to England and were trained in the art and technique of making optical glass. These physicists remained at our works for some two to three months and were attached to our own personnel, so that they had the fullest opportunity of learning all that we could teach them about the manufacture of optical glass. A number of other people were recruited in this country, some from the Canadian Forces here and some from our own staff, and were given a special intensive course before being sent to Canada for starting up the new process. Further, when manufacture began, one of our technicians went to Canada for some months in order to assist in overcoming the initial difficulties, so that the closest co-operation was maintained and is still being maintained. When it is realized that in November 1940 the optical-glass plant consisted of merely the walls of what was to be a factory, and that the first melting was successfully made in June 1941, and that by the end of 1941 a considerable number of meltings of first-quality glass had been made, it will be realized that, owing very largely to the enthusiasm and drive of the Canadian executives, a very remarkable result has been achieved. There has throughout been the closest co-operation between the Canadians and ourselves, and it can be said that there has been no serious hold-up or difficulty in transferring the manufacture from this country to that.

The factory * occupies about 16 acres, but only a portion of this is devoted to glass manufacture. The pots used are being made by a casting process which is essentially that employed in the Bureau of Standards, namely, the slip-cast method, by which clay treated with a small amount of alkali to give a fluid suspension is poured into a plaster mould. This method, which is well known, has the advantage of being more rapid than the old hand method, and also provides a pot which is considerably drier than that by the old process, and it therefore dries in a relatively short time (figure 1). When required for use, the pot is preheated in one of the batch of pot arches, then transferred to the gas-fired regenerative furnaces and filled with the appropriate batch. Because of the relatively small tonnage employed in an optical glass factory, a mechanical system of weighing is uneconomic, and, therefore, weighing and mixing are done by hand. The scoops and pans used for handling the batch are of stainless steel. All the sand before being used is run through a magnetic separator to avoid any risk of extraneous contamination. The city gas-fired furnaces are of the regenerative type and were built by American companies to R.E.L. designs. The empty

---

* For fuller details see *Glass Industry* (New York), **22**, Aug. 1941, p. 335.

pots are placed in the furnace by a hydraulic pot carriage, which goes under the name of "Winnie". The rest of the procedure for making, cutting up, moulding, examining, and so on, is an exact copy of that employed at our works in this country, and, in fact, the Canadian factory may be said to be a modernized version of our works. The differences are merely those which are possible because of the existence of an entirely new site and of the possibility of organization from the commencement. The types of glass being made at present are about six, which are the main types in general use for optical instruments. The clays available in Canada have not so far been found appropriate for the more extreme types of glass, but the manufacture of these is visualized, and the continuing co-operation between Canada and Britain will, when necessary, ensure the provision of pots of the appropriate resistance to attack to enable all types of glass to be made.

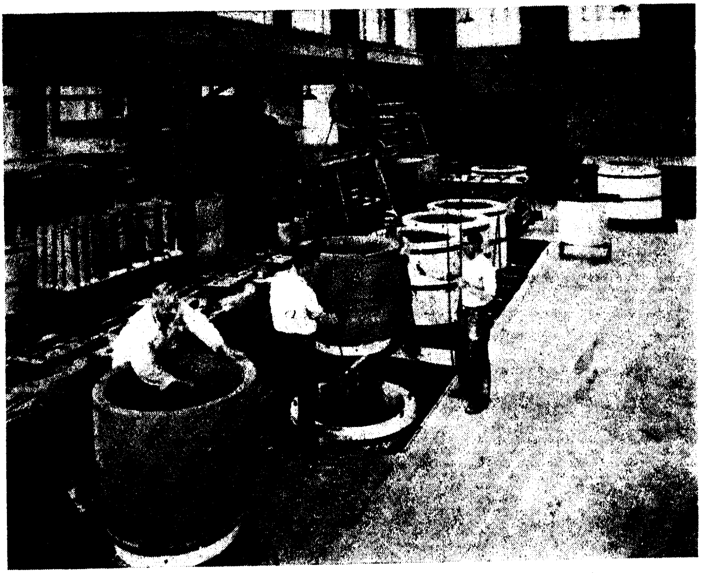


Figure 1. Slip-casting of pots at the Canadian factory.

Since the Canadian factory makes not only glass but also instruments, R.E.L.'s customer is in effect itself, whereas in this country the grinding and polishing work is done by separate firms, a method which has both advantages and disadvantages.

In considering the rapidity with which this works was put in hand, it should be remembered that at the time it was being built there was no lighting restriction, so that the factory was built continuously day and night, and there was at that time practically no restriction on the quantity of labour available. These comments are not intended to detract in any way from the achievement of Colonel Phillips and his colleagues, and I can only add that collaboration with them has been simple and complete.

There have been publications from time to time on developments in the manufacture of optical glass. During the last few years, greater developments have taken place than ever before, due to a sustained programme of research which has been carried out under the aegis of the Admiralty. It is clearly impossible for me to mention in detail many of the developments that have taken place, and I therefore find myself forced to pick out one or two items concerning which information has already been published in this country or elsewhere, and to deal with these only. It follows, therefore, that most of what I have to say today is not new in the sense that it has not previously been published, but it may be of interest to the members of this Society to try and present a connected picture of the various details that I have been able to select.

§ 3. RAW MATERIALS—SAND

Before the war, sand supplies for the manufacture of optical glass were derived practically exclusively from Germany. The sand deposits at Lippe and Hohenbockaer were of the highest quality known, and although substantial stocks had been accumulated or were existent in this country, a serious problem would have arisen had we not been able to find a native source. It had been known for some time (Turner, 1940) that in the Western Highlands at Loch Aline there existed a deposit of sand which appeared to be of fairly high purity as regards iron content. Owing, however, to the fact that it was in a very inaccessible position, the deposit had never been developed before. During the summer of 1940, Messrs. Charles Tennant & Sons, of London, in conjunction with Messrs. Charles Tennant & Co., Ltd., of Glasgow, undertook the development of the site, using the Union Oxide & Chemical Co., Ltd., as the sole agents for England. Considerable quantities have been sent to various glass manufacturers, apart from the requirements for optical glass manufacture. Although the rock on site is very hard, it breaks down readily to yield a granular material suitable in grain for glass melting. It will be realized that the first essential for the manufacture of optical glass is that the sand shall be as free as possible from colouring oxides, and particularly iron. The sand at Loch Aline contains about 0·02 % of Fe_2O_3, which is substantially higher than that of the foreign deposits. Water washing alone, however, reduces this iron content to about 0·012 % Fe_2O_3, which is comparable with Fontainebleau, but it is still substantially higher than the German sand, which used to run about 0·008 %. Intensive experimental work, both at Sheffield under Professor Turner and in our own laboratories, showed that it was possible by mechanical or chemical treatment to remove much of the iron-bearing material, and during the whole of 1941, optical glass in this country was made from Loch Aline sand which had been treated by a Wilfley mechanical separator, with the result that the iron content did not exceed 0·007 %, which is a lower average than had previously been achieved with the foreign material. Recently a new method of treatment has been developed, and sand is now available in sufficient quantities for optical-glass manufacture having an iron content not exceeding about 0·005 %. Thus at the present time the British optical-glass industry has a source of supply of sand which is of substantially higher quality than any which has previously been available. It should be noted, however, that for reasons stated above, and because of the cost of purification, the price of this sand is very substantially higher than that which was previously available. It is impossible to say at the moment whether it will be economic still to use this sand after the war, but the country is at least independent of foreign sources in case of emergency.

§4. HEAT TREATMENT AND REFRACTIVE INDEX

I will now turn to the question of recent research on the effect of heat treatment on the optical properties of glass, by summarizing briefly work carried out at the British Scientific Instrument Research Association, our own laboratories, and also the literature on the subject. There was an intensive period of research on annealing and its problems between the end of the last war and about 1926. This was all directed towards shortened time-schedule for annealing, and the sole criterion of success was the absence of double refraction. Glass at high temperatures is a viscous liquid, and there is no evidence of any sharp change to a definite solid state as it is cooled to room temperature. The viscosity-temperature curve is continuous, and at room temperature glass can be considered as a liquid of enormous viscosity. In cooling it from a high temperature, therefore, any temperature gradients which exist are, as it were, frozen into the glass, with the result that instead of being optically homogeneous and isotropic it becomes optically inhomogeneous and anisotropic, behaving as a uniaxial crystal. As a result of elastic deformation, variations in refractive index are present, and on examination in polarized light it is found to have different optical properties in different directions. Sufficiently slow cooling reduces these birefringence effects to a negligible amount, and the old test of an annealing schedule was that it should reduce the double refraction to something of the order of 0·01 λ per cm. It was, however, found on occasion that glass which appeared satisfactory from the point of view of double refraction and which was known to be free from striations still did not give a satisfactory optical image, and it was certainly possible to find two pieces of glass which, while appearing equally satisfactory in polarized light, gave substantially different pictures when examined on an interferometer (figure 2). At the same time it was clear that glass could be used satisfactorily under conditions where, had it been examined in a strain viewer, it would have been seen to be bad. The strain viewer is so sensitive to the effect of temperature changes that it is necessary to maintain a block at a constant temperature for a considerable time and to handle the block with cotton wool or other insulating materials during examination in order to get a satisfactory strain picture. In other words, the warmth of the hand in picking up a block is sufficient to change its appearance rapidly and seriously in the strain viewer, yet this same glass when used, say, in a submarine periscope which might rise suddenly from the temperature of the sea into 20° of frost or bright sunshine, gives an entirely satisfactory performance, so that it is clear its optical performance is not directly related to its strain-viewer appearance. Occasions have been known where lenses have been found unsatisfactory and have been returned for re-annealing, and although from the strain-viewer pattern they looked rather worse after re-annealing than they did before, the optical difficulties had disappeared and the glass was satisfactory. It is clear from all this that there is something other than double refraction which can be affected by the annealing process.

It has long been known that glass which is cooled rapidly has a lower refractive index than glass which has been cooled slowly from the same high temperature. These differences in refractive index are bulk differences, i.e. the same result is found whether the refractive index is measured at a surface, as in the Pulfrich refractometer, or whether it is measured through the bulk of the glass, as on a spectrometer. These differences are relatively large, amounting to, say, six or more units in the third decimal place. It can be shown by simple calculation that the plus and minus variations due to compression or tension in unannealed glass which is stressed to the tensile limit will not exceed more than one or two units in the fourth decimal place. These differences due to tension or compression are, therefore, almost completely masked by the general lowering of refractive index, corresponding to cooling rapidly from a high temperature. There has been a great deal of research on this question, as, for instance, on the effect of heating specimens from the same melting, some of which have been chilled and some of which have previously been annealed.

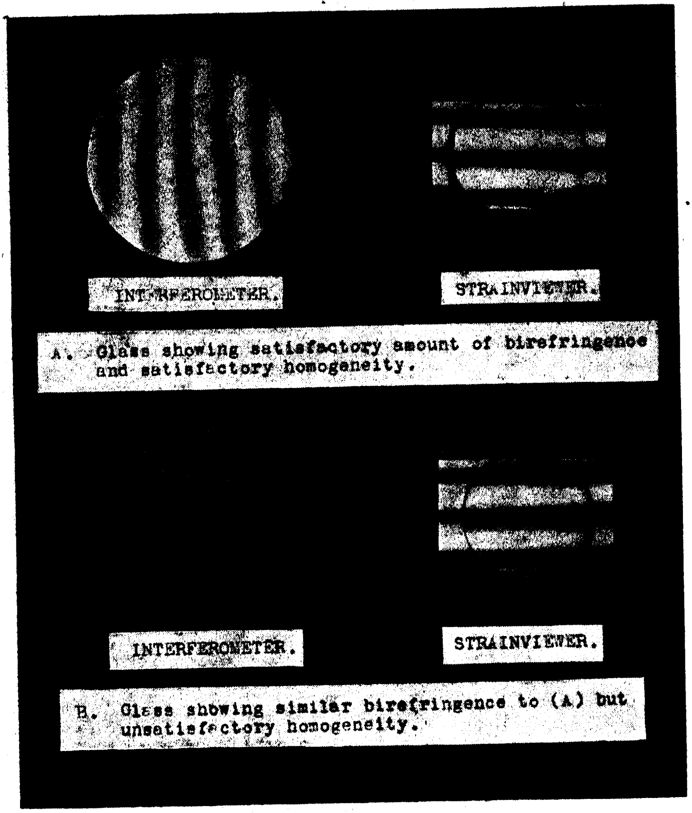

INTERFEROMETER.
STRAINVIEWER.
A. Glass showing satisfactory amount of birefringence and satisfactory homogeneity.
INTERFEROMETER.
STRAINVIEWER.
B. Glass showing similar birefringence to (A) but unsatisfactory homogeneity.

Figure 2. Interferometer and strain-viewer photographs.

Figure 3 shows the result of one such experiment (Lebedeff, 1926). Small specimens of the glass are heated at a constant rate and are extracted from the furnace at various temperatures. The pieces are allowed to cool rapidly in the air and the refractive index of the cold glass is measured. It will be seen that up to a temperature of some 500° or 600° c. there is no change in the refractive index of either sample. As the temperature rises, however, there is a marked change in the refractive index, which in the case of the well-annealed samples tends to fall until a temperature of about 800° c. is reached, beyond which it remains constant. In the case of the chilled specimens, the index rises a little and then, at about 800° c., falls to the same curve as the well-annealed glass.

A second series of experiments was to take well-annealed and chilled samples of the same glass and to maintain them at a constant temperature of, say, 550° c. (Twyman and Simeon, 1923; Winter-Klein, 1937). Specimens are extracted from the furnace as before and the refractive index of the cold glass measured. Figure 4 shows the results of such experiments. It will be seen that the well-annealed glass tends to fall in index and the unannealed glass tends to rise in index, and there are indications of an equilibrium index which is a function of the stationary temperature. It is also seen that the equilibrium index decreases as the temperature rises. By determination of equilibrium temperatures in this way, a curve of equilibrium index versus temperature can be plotted, and according to various authorities this curve is approximately linear over the range where, in the first experiments, the index was changing rapidly. The velocity with which the specimens of glass approach the equilibrium index is also a function of temperature, and this velocity increases very rapidly as the temperature rises. It is found that the rate of change of index with time at a constant temperature can be represented by a relation similar to that found by Adams and Williamson (1920) for the rate of change of double refraction. It was shown by the present author (Hampton, 1925/6) that the velocity constant was inversely proportional to the viscosity of the glass. Approximate calculations of the present velocity constant for refractive index, based on such meagre information as is available, suggests that a similar proportionality exists in the case of velocity constants for the approach of refractive index to the equilibrium index, and

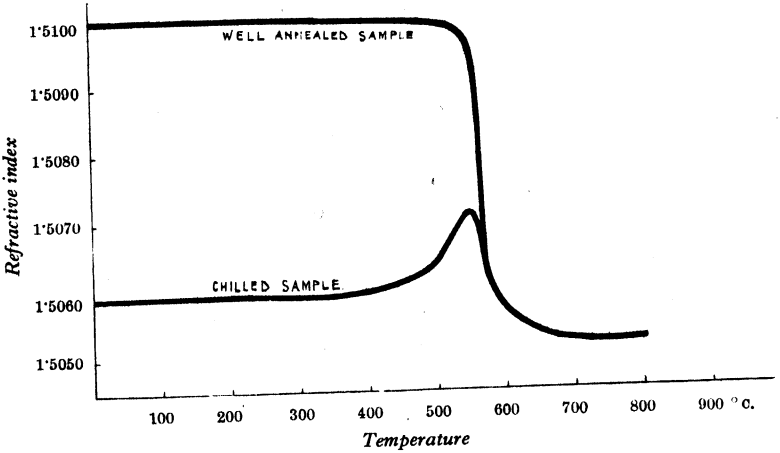


Figure 3. Effect of heating annealed and chilled samples of BSC glass.

although the matter cannot be considered as proved, it is almost certain that this velocity also is inversely proportional to the viscosity of the glass. Since the viscosity of glass falls very rapidly with increase of temperature, the velocity of approach also changes very rapidly, and may vary in the ratio 100 to 1 for a change of, say, 50° c. This fact makes the determination of the {equilibrium index, temperature} curve very difficult, and results at both ends of the annealing range are of doubtful value.

Without at the moment going into the question of what causes this change in the equilibrium index, it is clear that we have here an explanation of the curves in figure 4. Consider the well-annealed sample being heated to a high temperature and samples being removed from the furnace. At the lower temperatures the viscosity is so high that the velocity of approach to the equilibrium index remains what it was at room temperature. Above about 500° c. the viscosity of the glass is beginning to fall to a reasonable value and the refractive index of the glass turns towards the equilibrium-index line during the heating. It does, in fact, rather overshoot this line, and approaches it asymptotically. It will be noted that the experimental curve shows that the refractive index as measured on cold samples is constant above a certain temperature, whereas extrapolation of the {equilibrium index, temperature} curve suggests that the refractive index should continue to fall. It may be pointed out that the experimental evidence does not disprove the implication of the {equilibrium index, temperature} curve. The velocity of approach to equilibrium at these high temperature is so large that even during the high rate of cooling which is implied by the word " chilling " it is probable that the glass index can readjust itself to the equilibrium value corresponding to the temperature. Thus it is to be expected that the refractive index of a sample chilled from, say, 1000° c. will follow the equilibrium line for some distance until the viscosity becomes too

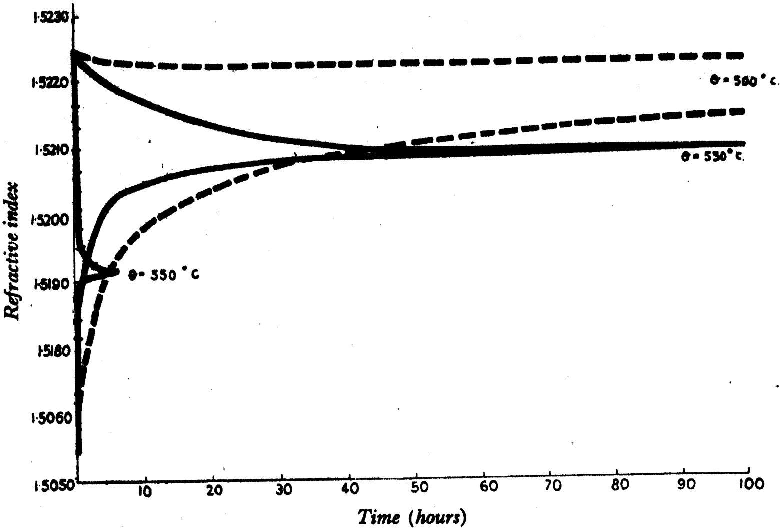


Figure 4. Effect of heating at constant temperatures on index of annealed and chilled samples.

high for this instantaneous adjustment to take place. The refractive index measured on the chilled sample, therefore, taken at 1000° c., would be recorded as the index corresponding to some lower temperature, and it follows also that, for a given rate of chilling, this recorded index will be independent of the temperature from which the sample was chilled, provided this temperature is sufficiently high. It may be shown experimentally that the limiting index is lowered with increase in the chilling rate.

Various experimenters, notably Lebedeff (1926) and Klein (1937) have assumed that these curves have indicated the existence of two forms of glass which, for want of a better term, they have called the α and β forms by analogy with the silica transformation. They say that glass below about 500° c. exists in the α form and above about 700° or 800° c. in the β form, and between those temperatures there is an equilibrium between the α and β forms. There are three facts which seem to me to throw grave doubt on this supposition:—(1) That the lower value of index reached can be shown to be a function of the rate of chilling during the experiment; (2) that a pure boric-oxide glass free from silica shows the same sort of effect (Turner, 1923); and (3) that the viscosity-temperature curve does not show any marked changes of direction as the temperature of the glass is raised.

It seems to me more reasonable to suppose that glass at very high temperatures consists of an intimate solution of simple silicates which on cooling tend to form aggregations into larger molecules. The important practical point is, of course, that the refractive index of a given glass depends on the effective temperature which is frozen in during the annealing operation. It is clear that the slower the cooling, the higher will be the refractive index.

These results have an important bearing on the annealing of slabs of glass. If a piece of glass is uniform in temperature while being held prior to the annealing operation, the only index changes introduced into it will be the plus or minus variations due to the temperature gradient during cooling, and these will probably be small, perhaps not more than a few units in the fifth decimal place. If a small piece of glass is held at, say, 550° c. until the index has approached the equilibrium and is then cooled slowly from this temperature, it will have a general refractive index slightly higher than the equilibrium index for 550° c., together with the plus or minus variations due to the gradient. If it is held at some other temperature, say 540° c., it will have a final index slightly higher than the equilibrium index for 540° c., together with the plus or minus variations, but the two specimens of glass will have different refractive indices, although their double-refraction pattern may be the same. If, therefore, both these temperatures exist during the annealing operation, that is, if a temperature difference exists during the soaking period, then different parts of the finished plate will have different refractive indices. Without the most elaborate precautions it is very difficult to get any sort of annealing kiln where the temperature differences are so small as not to give a risk of departures across a slab of optical glass of the order of 10° c. Much experimental work, therefore, has been directed towards the provision of annealing equipment where the temperature differences are negligibly small, and kilns have been designed and are in use where the temperature difference over a length of 2 feet or more does not exceed 3 or 4° c. These kilns are capable of being cooled at a much higher rate than the old type, while still yielding glass which is of the highest standard when viewed on the interferometer, although the strain-viewer pattern may not be appreciably different from that given by less uniform kilns. Experience has shown that the glass which is satisfactory on the interferometer, due to the absolute uniformity of index, is capable of being used for the very highest types of optical instruments, and it is becoming generally accepted in the optical industry that the interferometer, and not the strain-viewer, is the instrument to be used. It is now a routine operation for samples from all kinds of optical glass to be examined on the interferometer, and for the strain viewer to be used to a considerably less extent.

There was for many years an objection, which was based on painful experience, that the highest quality prisms could only be obtained by sawing from large blocks, and not by moulding. Since moulding is much more economical in the use of optical glass than cutting up large pieces, recent developments in annealing, which have proved beyond doubt that moulded prisms can give results equal in quality to those produced by any other method, have been of the utmost importance in amplifying the supply under the present emergency conditions.

§ 5. DURABILITY

Another question which I would like to discuss briefly is that of the determination of the durability of optical glass. The object of such a test is to obtain a measure of the resistance of a polished surface of glass to atmospheric agents, since this surface in an optical element is subject in use to exposure to the varying conditions of humidity and the possible chemical attack of atmospheric impurities. The varying humidity of the atmosphere results in corresponding variations in the water content of the glass surface, accompanied by alternations of deposition and evaporation of moisture. These alternations result in the separation of soluble constituents from the glass and the production of a surface film. Compared with its optical properties, the durability of optical glass is of secondary importance, but it is essential that in obtaining any desired optical properties, either by developing a new type of glass or modifying an existing type, the final choice of batch should be decided by the durability. It is obviously important, therefore, that the manufacturer should be able to determine fairly rapidly the effect on durability of variations in batch composition, and it is also desirable that the user should be furnished with some indication of the probable performance under service conditions. Whereas with well established types, observations on glasses which have been submitted to prolonged exposure to atmospheric conditions are fairly well known, such a method of testing is clearly impracticable as a means of forecasting what a new glass will do. In addition, owing to variations in conditions during actual exposure, such information as is available is of a statistical character. Several laboratory methods have been put forward for determining the durability, but such accelerated attacks as a rule differ markedly from those to which the glass is exposed in use. The results obtained by such methods are sensitive to slight variations in experimental conditions, and in consequence often fail to give agreement even among themselves. In addition, in the few cases where comparison has been possible with the results obtained under natural but severe conditions, there appears to be little correlation. The tests which have been employed and discussed in the literature are the following:—

Chemical resistance tests (see, for example, Taylor, 1936)

This method is used to determine the chemical resistance of glasses to chemical reagents. The glass is usually in the form of a powder graded between moderately close limits and the solubility of the powder in the reagents is determined. This test is very drastic in character compared with the weak chemical action of the atmosphere, and in general there is little relation between the results given by it and those found in use.

Autoclave test (see, for example, Baillie, 1922)

In this test a polished glass surface is subjected to the action of superheated steam. The rate of corrosion is far more serious than is met with in use, and the conditions of humidity to which the glass is subjected are quite different from those in practice. The test does not as a rule include the variation of temperature conditions, which forms one of the most prominent features in practical use.

Dimming test (see, for example, Elsden, Roberts and Jones, 1919)

This consists in subjecting the surfaces of the glass to the action of air saturated with water vapour at some uniform temperature, usually 80° c., for a definite period. This test does not allow for the humidity variation already mentioned, and very long periods are necessary in order to distinguish between good glasses.

Iodoeosin test (see Mylius, 1910 and 1913)

Results based on the Mylius iodoeosin test are quoted in certain optical-glass catalogues to specify the durability of the various types. A quantitative measurement is usually made of the amount of free alkali in a freshly fractured surface which is available for combination with iodoeosin. The surface is then subjected to exposure to a moist atmosphere at 18° c. for seven days, and the change in the amount of free alkali is determined. Thus, increase in the amount of free alkali is taken as lack of stability, and decrease of free alkali is associated with stability. While it is accepted that the test is probably reliable for glasses containing appreciable amounts of alkali, it is misleading for those types containing lead oxide, zinc oxide and boric oxide.

Electrical test (see B.I.R.A. reports R. 19 and R. 26)

The British Scientific Instrument Research Association devised a test for the durability of optical glass surfaces in which the electrical conductivity and changes of conductivity under accelerated weathering conditions are measured. As with the iodoeosin test, this is only reliable for certain types.

Some recent work on the durability of optical glass has been published in America (Jones, 1941). Observations were made on the behaviour of polished glass surfaces exposed to ordinary service conditions, and two types of surface change due to weathering were noted. Glasses which had been subjected to a humid atmosphere became covered with a hazy film of soluble alkaline salts, and this process was called " dimming ". On the other hand, contact with condensed moisture led to the separation of lead or barium oxides and the formation of a silica-rich film. The film exhibited interference colours and the process was called " staining ". Various accelerated durability tests were investigated, and it was concluded that two tests were required: one to determine the tendency to dimming and the other to determine the tendency to staining. For the dimming test, exposure to a saturated atmosphere for 28 days at a constant temperature of 50° c. was employed. The best test for staining was found to be that developed by Berger (1936) using nitric-acid solution of known strength at 25° c. and observing the time required for the formation of a surface film showing a dark blue interference colour.

The test which I propose to describe now was developed in this country over a period of some years, and, in fact, was being used as a routine method before the recent American work was published.

In attempting to devise a new test for the determination of durability it was laid down that the method should be capable of discriminating between slight differences in durability, and the variables should be capable of accurate control; it was also considered essential that the results should be comparable with those that are obtained under natural conditions by exposing specimens of the same glasses to the action of the atmosphere in the tropics for periods of one to two years. In practice, the instruments are in many cases subjected to a relatively high temperature during daylight and a low temperature during the night. The result is that moisture condenses on the glass at night and is evaporated off during the day. An apparatus was therefore constructed in which cycles of temperature were introduced, and the name *thermodyne* was coined to describe it, in contradistinction to the thermostat.

The thermodyne method of determining the durability of glass surfaces consists essentially of submitting either freshly broken or optically polished surfaces (sealed off in a flask together with a quantity of water) to a series of temperature cycles from room temperature to 60° c. in an atmosphere of air saturated with water vapour for a period of 12 days. Each cycle is of two hours' duration. The attacked surfaces are subsequently examined visually and photographed for record purposes. The type of attack which results is characteristic of the type of glass. The degree of attack is recorded by an arbitrary number, comparison always being made with the attack on specimens of similar standard type tested simultaneously. Control is possible on the following points:—(*a*) Upper and lower temperature limits and the duration of the cycle; (*b*) humidity of the atmosphere in contact with the specimens; (*c*) composition of the atmosphere in contact with the specimens; (*d*) condition of the test surfaces.

The results obtained with the test were compared systematically with the results which had been obtained by exposure of the same glasses to tropical atmospheres. Without going into details of the experimental procedure, it may be said that this method of test has given valuable results.

It may be pointed out that both the conditions for producing the two types of weathering noted by the American experimenters are embodied in our method, and while it is recognized that no single test is available for determining the durability of the wide range of optical glasses manufactured, comparison of the results of the thermodyne test with the results of natural-exposure tests shows that the accelerated test is fairly reliable.

In typical tests on binocular prisms returned from service, the specimens appear in the same relative order when ranged according to their state after this test or after natural exposure; moreover, the two specimens which were badly stained naturally show heavy staining in the accelerated test.

It may be said generally that the thermodyne attack is not exactly the same as the natural attack. In fact, the thermodyne attack is rather more severe than that resulting from years of exposure in the tropics, but the more severe nature of the accelerated attack accentuates differences in durability. At the same time the test is not so severe as some of those already described, and it is perhaps most valuable in assessing the durability when modifications of an old type of glass are being developed.

It is possible that with the development of non-reflecting surface films, this question of glass durability will become of less importance, and it is worth noting that Jones and Homer (1941) have recently used the method employed for their accelerated weathering test to produce silica-rich films which have later been stabilized by baking to form a durable film.

§6. CONCLUSION

I do not make any suggestion that the four problems I have dealt with are the only ones with which we are concerned, nor even that they are the most important ones; they were simply taken as a means of indicating to you the type of problem that has to be met in the manufacture of optical glass. I feel that a Society such as this will give opportunities for closer collaboration between makers and users of glass, and I would like to point out that while there are many lines on which development of new types of optical glass may proceed, it is only in a few of these that any useful result is likely to be achieved. In order to decide which of these calls for the closest collaboration between lens designers and the makers of optical glass, it would be of interest and value to us as manufacturers if the representatives of the various lens-designing and lens-making firms could give fairly precise indications as to the type of {refractive index, dispersion} relation that would be of most value to them in designing new systems. At any time, there is a definite restriction on the amount of research which can be undertaken, and we do not wish to extend our exploration over parts of the problem which do not seem likely to lead to fairly rapid results. If, therefore, either through this Society or otherwise, it could be so arranged that the users could give precise figures of index and dispersion that would be of value, much unnecessary work could be avoided. Certain optical firms have already indicated to us the types of glass which they think will be needed in the future, and we hope that other firms will follow suit.

REFERENCES

ADAMS, L. and WILLIAMSON, E., 1920. *J. Franklin Inst.* **190**, 597.

BAILLIE, W. L., 1922. *J. Soc. Glass Tech.* **6**, 279.

BERGER, E., 1936. *J. Soc. Glass Tech.* **20**, 257.

ELSDEN, A. V., ROBERTS, D. and JONES, H. S., 1919. *J. Soc. Glass Tech.* **3**, 52.

HAMPTON, W. M., 1925–6. *Trans. Opt. Soc.* **27**, 163.

JONES, F. L., 1941. *J. Amer. Cer. Soc.* **24**, 119.

JONES, F. L. and HOMER, H. G., 1941. *J. Opt. Soc. Amer.* **31**, 34.

LEBEDEFF, A. A., 1926. *Rev. d'Optique*, no. 1, 1.

MYLIUS, F., 1913. *Silikat Z.* **1**, 2, 25, 45 ; 1910. *Z. Anorg. Chem.* **67**, 200.

TAYLOR, W. C., 1936. *J. Soc. Glass Tech.* **20**, 405.

TURNER, W. E. S., 1923. *J. Soc. Glass Tech.* **7**, 139.

TURNER, W. E. S., 1940. *J. Soc. Glass Tech.* **24**, 197.

TWYMAN, F. and SIMEON, F., 1923. *J. Soc. Glass Tech.* **7**, 199.

WINTER-KLEIN, N., 1937. *Rev. d'Optique*, no. 11, 361.

DISCUSSION

DR. SPENCER JONES. I am very interested in what Dr. Hampton has said about the thermodyne test for dimming. In the last war I was much concerned with the trouble from dimming of glass surfaces in optical instruments used under tropical conditions. I used both the autoclave and iodoeosin tests and found that there was a close correlation between the results of these tests and the liability to dimming in hot moist climates, though, as Dr. Hampton has mentioned, the tests were not in every respect satisfactory. It was found that, for the same type of glass, the trouble from dimming was dependent both upon the perfection of polish of the surface and upon the care taken in cleaning it. Dr. Hampton's tests are made mainly upon freshly fractured unpolished surfaces, and I should like to ask him whether the thermodyne tests show that the dimming is dependent both upon the perfection of polish of the surface and upon whether or not it is chemically clean.

AUTHOR'S reply. It should be pointed out that with the thermodyne test, it is possible to determine the liability to staining as well as the liability to dimming. Both the autoclave and the iodoeosin tests give misleading results for glasses which stain badly under natural conditions, and our experience with the thermodyne test shows that it gives more reliable results for such glasses, particularly for dense barium crowns.

With regard to the condition of the test surfaces, a considerable amount of work was done on this subject during the development of the test. It was found, for example, that unless the surfaces were chemically clean, there was considerable variation in the results. The criterion of cleanliness adopted for the test is that the surface can be uniformly wetted by distilled water.

It has also been found that for certain types of glass the results depend on the perfection of polish of the test surface. Thus, surfaces polished by an inexperienced polisher appeared to be much less durable than those polished by an experienced polisher, and the results for freshly fractured surfaces agreed with those obtained for surfaces polished by the experienced polisher. This result is in agreement with the conclusions reached by Lord Rayleigh (*Proc. Roy. Soc.* A, 1937, **160**, 507) that the refractive index of a polished glass surface depends on the method of polishing, and that for commercial polishing the refractive index of the polished surface approximates to that of the body of the glass.

For this reason, freshly fractured surfaces are used whenever possible in the thermodyne test, and in such a case the only cleansing necessary is the removal of glass débris which arises from the fracture.

THE MEASUREMENT OF RADIATION FOR MEDICAL PURPOSES

By W. V. MAYNEORD,
The Royal Cancer Hospital (Free)

Lecture delivered 27 *February* 1942

ABSTRACT. The measurement of radiation for medical purposes is primarily concerned with the measurement of absorbed energy either at a given element of mass within the irradiated tissues or throughout the body as a whole.

The ionization method of measuring x rays and gamma rays is briefly described in the paper and the fundamental nature of the röntgen is discussed. The distribution of radiation throughout a body subjected to x or gamma irradiation is next considered, and examples given of the variation of dosage-rate in typical cases along the axis of a beam sent into soft tissues. This leads to a study of " isodose surfaces " and some simple geometrical methods of obtaining them.

The relationships between the dose in röntgens and the energy absorbed per unit mass of material or the energy flux per square centimetre of cross-section of the beam are indicated.

The second problem of total energy absorbed is then attacked by integrating the product of the dose and element of mass throughout the whole mass of interest, thus obtaining the so-called " integral dose ". Again, simple geometrical and analytical methods of deducing the values of the quantities involved are described. The significance of total energy absorption in many problems (as, for example, that of protection against stray radiation) are indicated. Finally, the total amounts of energy involved in the therapeutic use of ultra-violet and infra-red radiations are compared with those of x rays and gamma rays, and a plea is entered for the correlation of all these radiations on an energy basis.

§ 1. INTRODUCTION

THE methods of measurement of radiation for any purpose whatsoever must in the end be determined by the fact that we are concerned with a form of energy. Nevertheless the specific purpose to which a radiation is to be put often decides which of a number of quantities it is of greatest importance to investigate. Left to a severely physical investigation of a beam of x rays or gamma rays, we might be tempted to concentrate on the measurement of its intensity, but if the beam were intended for use in radiation therapy our efforts would then be to some extent misdirected.

The whole of the study of the biological and medical use of radiation is dominated by the rule that absorbed energy alone is effective, so that we are forced in considering the physical aspects of the use of radiations for these purposes to look very closely into the measurement of this quantity. This study of absorbed energy, either at any point in the tissues, or throughout the body as a whole, forms the subject matter of this paper.

We may commence with a very short survey of the problem in relation to the ionization system of measurement of x or gamma rays in common use and embodied in the unit known as the röntgen. This leads to an attempt to estimate the total energy flowing into the body or absorbed by it from beams of x and gamma rays, and eventually to link these values to measurements of ultra-violet and infra-red radiations also used for therapeutic purposes. Our survey is often rough and inexact, but the time seems ripe for an attempt to sketch out the ground, perhaps in this way indicating where detailed examination and experiment may now be most profitable.

§ 2. THE IONIZATION METHOD OF MEASUREMENT OF X RAYS

Among the various physical or chemical effects of x rays, the ionization produced by them in air stands out pre-eminently as the most promising basis on which to construct a system of measurement for medical use. This suitability depends upon many factors, three of which may be noted. First, the measurement of an ionization current may be carried out with fair accuracy over a very wide range of values. Indeed, the measurement itself is often easy compared with the decision under what conditions to measure and what the current means when we have measured it. Secondly, the energy required to produce a pair of ions in air is nearly independent of the speed of the electrons producing it (Eisl, 1929), and so independent of the wave-length of radiation setting the electron in motion. This, together with the fact that the effective atomic number of air is often approximately the same as that of soft animal tissues (Mayneord, 1938), means that the ionization observed in air may be made a suitable basis for the estimation of the energy absorbed by tissues over a wide wave-length region stretching from perhaps 0·01 to 2·0 A. This so-called " independence of wave-length " is very valuable in medical radiology, where the rival, and often conflicting, claims about the superiority of different wave-lengths are keenly debated.

The conditions of measurement thought to be necessary and sufficient to obtain significant ionization currents may be seen in the original definition of the röntgen as adopted by the International Congress of Radiology in 1928.

§ 3. DEFINITION OF THE RÖNTGEN (1928)

" The röntgen is the quantity of x radiation which, when the secondary electrons are fully utilized and the wall effect of the chamber is avoided, produces in 1 c.c. of atmospheric air at 0° c. and 76 cm. of mercury pressure such a degree of conductivity that 1 electrostatic unit of charge is measured at saturation current."

Immediately following the adoption of this unit, experimental and theoretical work was commenced which was intended to extend its use to x rays produced at very high voltages and to gamma rays, since, apart from the theoretical interest of the problem, this was a pressing everyday need. At first considerable difficulty was encountered owing principally to the very great ranges of the secondary electrons set free in air at normal pressures, and the difficulties of eliminating secondary radiations, but eventually the results of measurements with air-wall and parallel-plate chambers were reconciled, and the problem now seems to be fairly satisfactorily solved to the accuracy required for medical use (Gray, 1937; Mayneord and Roberts, 1937; Kaye, Binks and Perry, 1939). It may be added that the research often appeared as the determination of the dose in röntgens received in one hour at a point one centimetre from a point source of radium filtered with 0·5 mm. platinum. Our own determination of this quantity with air-wall chambers led to the value 8·3 r., in a fair agreement with more recent results obtained by Taylor and Singer (1940) in measurements with a high-pressure parallel-plate chamber which give 8·16 r. as the final value, and probably represent the most accurate determination of this constant to date.

It was not until 1937 that the röntgen was provisionally adopted for gamma rays as well as x rays, a revised definition being at the same time promulgated. This new definition runs as follows:—" The röntgen shall be the quantity of x or gamma radiation such that the associated corpuscular emission per 0·001293 gramme of air produces, in air, ions carrying 1 electrostatic unit of quantity of electricity of either sign."

This new definition aimed at improving the rather clumsy form of the original, extending its use to gamma rays, and also at eliminating from that definition the statement of experimental conditions which gave it the form of the description of a standard rather than that of the definition of a unit. The fundamental idea, namely, that we take all the electrons set free by the radiation from a certain mass of air, utilize as much as possible of their energy in producing ionization in air and measure the number of electrostatic units of charge thus liberated, appears much more clearly in the new definition.

§ 4. DISTRIBUTION MEASUREMENTS

We now turn to a second aspect of the measurement of x or gamma rays in medicine, namely, the results of measurement throughout a scattering mass. These measurements are undertaken so as to find out how the energy sent into such a mass is distributed, and so, if we can, deduce the energy absorbed per unit mass at each point throughout the absorber, believing that this quantity is of fundamental importance in determining the biological effects of the radiation at each point.

The measurements are normally carried out by substituting for the human body a tank of water throughout which measurements of dose or dosage-rate (that is röntgens per unit time) are made. Owing to the wide range of radiations of interest, different experimental arrangements have naturally to be employed, for we have to solve the problem at low kilovoltages for radiations which penetrate little more than a fraction of a millimetre, while for very high exciting potentials we have radiations which only fall to half their surface dosage-rate at a depth of some 10 cm. to 15 cm. Clearly very different ionization chambers and accessories are required in these two cases.

As an illustration of the kind of results obtained in the normal therapeutic regions of 200 kv. to 2000 kv., we may make use of a survey (Mayneord and Lamerton, 1941) which sets out some of this information.

Considering first of all the change in dosage-rate as we enter the medium along the axis of the beam sent perpendicularly into it, we find that the apparent penetration depends upon three main factors, namely, average wave-length of the radiation; focal distance, that is the distance from source of x rays to the surface of the absorber, and finally to a very great extent on the area of cross-section of the beam which in its turn fixes the mass or volume from which scattered radiation is received by a point on the axis. These factors were analysed in our survey and tables drawn up from which the doses at a depth may be read for a large number of conditions. Typical examples of the penetrations observed are given in figure 1. Detailed mathematical analysis of the problem has not yet been successfully undertaken, though by the use of the Klein-Nishina scattering formulae, coupled with simplifying assumptions, the main features may be accounted for. The problem is, however, a very complex and difficult one.

Of more interest to us at the moment is the complete distribution of radiation in any plane containing the axis of the beam, but of course the experimental measurement is now much more prolonged, so that correspondingly less information is available. However, during the last few years data have been collected over a very wide range of conditions varying from 45 kv. and 2 cm. focal skin distance (Lamerton, 1940) to the gamma rays of radium filtered with 1 cm. of lead and used at a focal distance of 75 cm. (Mayneord, 1933; Honeyburne, Lamerton, Smithers and Mayneord, 1939). In particular, in the region of 200 kv. to 400 kv., many measurements have been made, and we may make use of some of these in what follows.

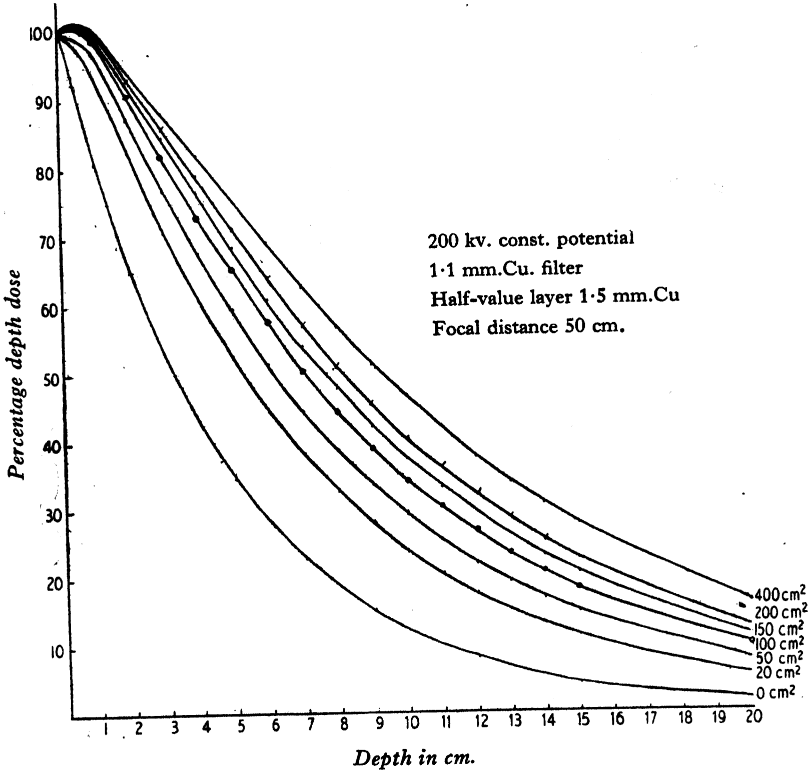


Figure 1. Variation of the dose with depth for a fixed quality of radiation but different field sizes.

Consider, for example, the results shown in figure 2. These show the distribution of radiation in a plane containing the central axis of a circular beam of x rays sent perpendicularly into a mass of water. The curves are drawn joining places having the same dose, and are, therefore, known as *isodose curves* or *dose contours*, all numbers representing the dose at a given point expressed as a percentage of the dose recorded at the centre of the beam on the surface. This quantity is usually known as *percentage depth dose.* Many hundreds of these sets of curves have been drawn from experimental observations and are of the utmost practical value in that, as transparencies, they may be laid on cross-sectional diagrams of the body and the dose received at a depth may be estimated for any conditions of irradiation or arrangement of converging beams which are proposed. Modern radiation therapy is increasingly dependent on such distribution diagrams.

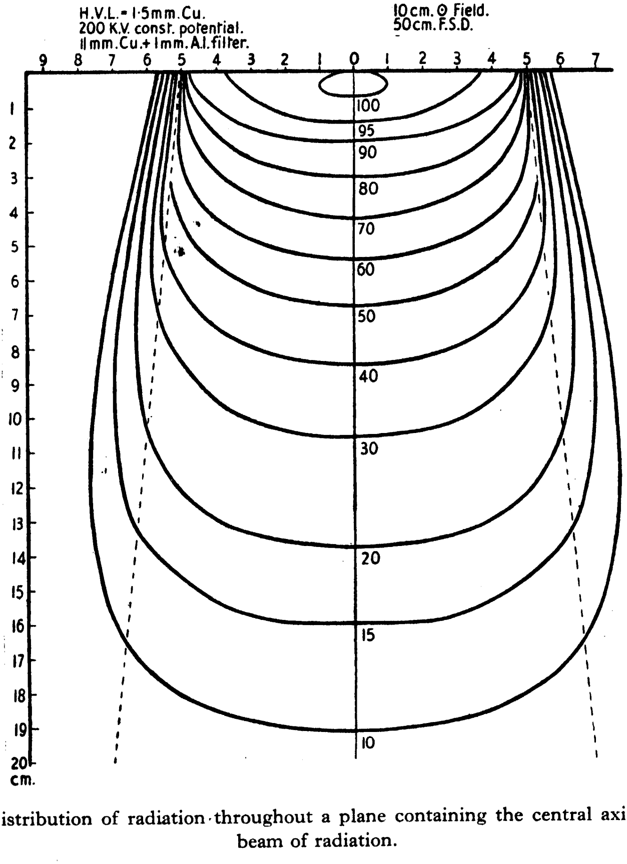


Figure 2. Distribution of radiation throughout a plane containing the central axis of a circular beam of radiation.

Our task is, however, only just beginning, for it is found that the problem as set for the hospital physicist is usually a three-dimensional one of great complexity in which beams of radiation are sent in at all kinds of angles to each other, while the axes of the beams are by no means invariably confined to one plane. In order to solve the three-dimensional problems, the most promising method of attack seems to be to develop first of all methods of finding the distribution of radiation in any plane however inclined to an incident beam. If the beam is circular, then clearly, on rotation of any distribution diagram such as figure 2 about its axis, each of the isodose curves generates an isodose surface. The determination of the distribution of radiation in any plane, however inclined to the axis of the beam, is now seen to be the determination of the intersections of these isodose surfaces by that plane. A simple geometrical device (figure 3) may be used to determine the required intersections (Honeyburne *et al.*, 1939). The diagram of the radiation of interest may be drawn on glass or Perspex so as to be capable of rotation in a yoke CDC′, while at the same time the axis of the beam AB may be set at any inclination to, and at any height above, a drawing board G on which the required diagram is to be traced. A vertical steel pointer J, ending below in a sharp pencil K, is adjusted to the height of the plane in which the distribution is desired. If the pointer J be brought into contact with an isodose curve of the diagram, then on depressing the pencil, a mark is made at a point vertically below a point in the intersection of the isodose surface and the plane in which the pointer moves. By simultaneous rotation of the diagram and adjustment of the pointer of the marker to a given isodose curve, the shape of the intersection of the plane with this isodose surface is rapidly traced. In this way the distribution of radiation in any plane may be quickly found, and by studying a number of parallel planes and drawing the diagrams in coloured ink on glass very instructive transparent models of the isodose surfaces may be constructed (Honeyburne *et al.*, 1939).

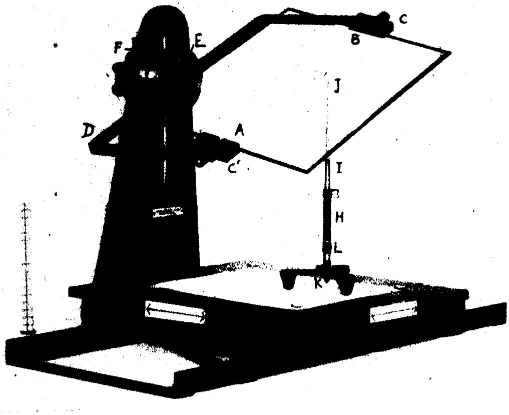


Figure 3.

We have found this contour projector a useful instrument for such rather academic exercises, but for solving the practical problems set by our medical colleagues it is still inadequate. A more powerful weapon is the dose-finder, which is a development of the contour projector.

The dose-finder (figure 4) in its present form (Honeyburne *et al.*, 1939; Mayneord and Honeyburne, 1941) consists of three accurate slides A, at right angles to each other, carrying a number (usually four) of steel pointers of adjustable length. Suppose now the distribution of radiation is required within a patient to be subjected to several beams: we make a plaster mould (B) of the patient and set up this mould on an adjustable stand so that one of the pointers may be adjusted to reach all points of interest within the mould. A model of the end of the applicator (C) defining the beam is then set in relation to the mould exactly as the real beam will be adjusted on to the patient. The mould is then removed and set in the same relation to the second pointer as formerly to the first. Meanwhile a distribution diagram D of the first beam is slipped into the applicator so as to be capable of rotation. A second model E of the applicator of a possible second beam is now similarly adjusted into position on to the mould (as shown in figure 4) and the process repeated. This process is continued until all the beams are correctly set each to its own pointer. Since all the pointers

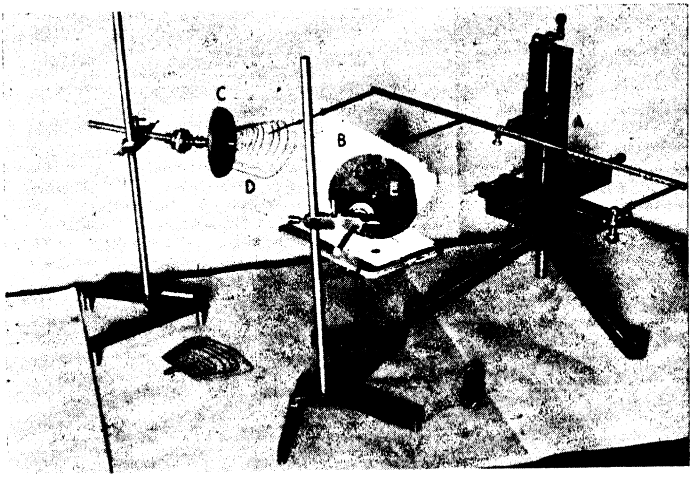


Figure 4. Dose finder. The instrument is shown at a time when a second beam of radiation is being adjusted to a model of the patient.

correspond to the same position in space, to estimate the total dose received we rotate the distribution diagrams until they touch their respective pointers, read off the various doses, and by simple addition (preferably with an electric adding machine) find the total at that point. By movement of the slides the total dose at any position in space may be found, any plane of interest investigated, or any other studies carried out. As an interesting accessory we may use a skeleton or a set of anatomical cross-sections set in relation to yet another pointer and so transform our Cartesian coordinates into anatomical location, to the great relief of our medical colleagues.

The details of the medical applications of these instruments do not concern us here, but we may add that they are equally applicable to the study of radium techniques using teleradium apparatus or the distributions around groups of radium needles, or, indeed, any three-dimensional problems of these kinds (Mayneord and Honeyburne, 1941). The descriptions are merely inserted here to show how the isodose surfaces may be studied, and in the hope that the instruments may be of interest to physicists faced with similar problems in other fields.

By these methods we may make an estimate of the dose at any point throughout a medium traversed by these radiations, and we now turn to another aspect of the problem, namely, that of the energy absorbed both at a given point and also throughout the medium as a whole. We must therefore now consider the relationship between the röntgen and energy absorption.

§ 5. THE RELATIONSHIP OF THE RÖNTGEN TO ENERGY ABSORPTION

It will be recognized that the röntgen is essentially a measure of energy absorption in air, the factor of proportionality depending upon the energy required to produce one ion pair. This quantity has been extensively investigated (Gray, 1936), and it seems that its value is approximately 33 electron-volts. It follows that a dose of one röntgen corresponds to the absorption of 0·11 erg per 0·001293 gramme of air, or approximately 85 ergs per gramme of air per röntgen. The energy absorption per röntgen in other media may be deduced from theoretical considerations first set out by W. H. Bragg in 1912, who showed that the energy absorption in a medium traversed by penetrating radiation depends upon the real energy absorption coefficient of the medium and the range of the secondary electrons in it. Recent studies by Gray (1937) have placed the theory upon a more exact basis. We may therefore calculate the energy absorption per electron in a number of media of biological interest (Mayneord, 1938), whence it appears that for wave-lengths below 0·1 A., materials of low atomic number are practically identical, but in the region of 0·1 A. to 0·3 A. rapid variation may occur. Beyond 0·3 A., the relative values depend principally on the fluorescent absorption coefficients, and are, therefore, approximately independent of wave-length. It seems that " soft tissues " and " air " are similar to each other, and, therefore, the energy conversion in air is an approximate measure of the energy absorption in soft tissues provided there is electronic equilibrium.

§ 6. RELATIONSHIP BETWEEN INTENSITY AND DOSAGE-RATE

Consider a parallel beam of radiation incident upon an element of mass *dm* of air. Then if I is the intensity of the beam in ergs/cm.² sec., the energy converted per second into electronic motion is given by

$$I(\sigma_a+\tau)dm\,.\,N, \qquad \ldots\ldots(1)$$

where N is the number of electrons per unit mass of material and σ_a and τ are respectively the real absorption coefficients per electron associated with recoil and photoelectrons. We have seen, however, that a dose of one röntgen delivered to a mass *dm* of air corresponds to energy absorption 85 *dm* ergs, so that if D_s is the dosage-rate in röntgens per second,

$$D_s = I\frac{(\sigma_a+\tau)\,.\,N}{85}\ \text{r./sec.} \qquad \ldots\ldots(2)$$

Alternatively, if total energy E_0 flows per unit area in a time t, then the total dose recorded (D) is given by

$$D = E_0(\sigma_a + \tau)\frac{N}{85}\text{ röntgens.} \qquad \ldots\ldots(3)$$

It is of considerable interest to calculate the total flow of energy per square centimetre per röntgen from this equation (Mayneord, 1940 a), the result being shown in figure 5. Some of the values given are similar to those of Lauritsen (1933), but cover a wider range of wave-lengths and are displayed rather differently.

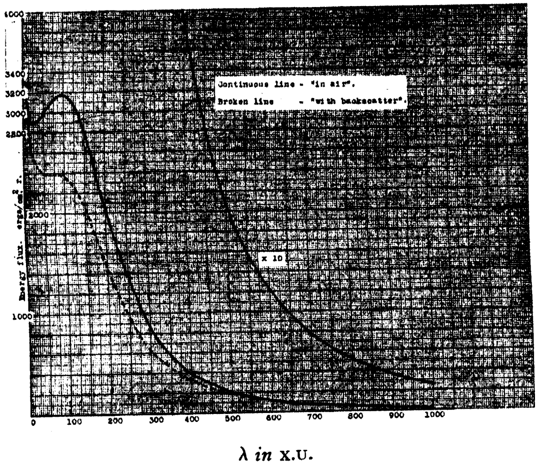


Figure 5. Energy flux per röntgen at different wave-lengths.

It will be seen that in the short wave-length region, the total energy flow per röntgen is of the order of 3000 ergs/cm², but falls to very low values (27 at 1 A.) for long wave-lengths. This distinction between dosage-rate and intensity is of great theoretical and practical importance.

§ 7. TOTAL ENERGY ABSORPTION

Consider now the parallel beam of radiation of cross-section A cm² and of intensity I ergs/cm² sec. entering a mass of absorbing material similar to air for a time t. Imagine, too, that we make distribution measurements throughout the absorbing mass and find for every element of mass dm the dose D_m in röntgens delivered to it. Let ρ be the electronic density of the medium compared with that of air. Then the energy absorbed throughout the scattering mass may be written

$$\text{Total energy absorbed} = \kappa\rho\int^{m} D_m \, . \, dm, \qquad \ldots\ldots(4)$$

where $\kappa\rho$ is a constant of the order of 85 ergs/gm. röntgen and the integral is taken throughout the absorbing mass. Clearly, if all the energy were absorbed we would expect

$$AIt = \kappa\rho\int^{m} D_m \, . \, dm. \qquad \ldots\ldots(5)$$

We see, therefore, that the distribution studies already described contain a clue to total energy absorbed in a scattering mass, and we turn, therefore, to the various methods of finding the integral of equation (4). It is convenient to take the energy absorption when one röntgen is delivered to one gramme of air as a unit. We shall refer to it as one gramme-röntgen. The quantity obtained by integrating the dose throughout a given mass we shall refer to as the *integral dose* throughout that region. For medical purposes it is found that a gramme-röntgen is far too small a unit, since we irradiate thousands of grammes with doses of thousands of röntgens, so that one million gramme-röntgens (a megagramme-röntgen) is found to be of a very convenient order of magnitude. If we have a medium of density unity and atomic number equal to that of air, one megagramme-röntgen is approximately equal to 2·02 gm. calories, a convenient conversion factor when estimating total energy absorbed in heat units.

The method given below is the logical extension of the methods of estimating radiation energy from ionization produced in media surrounding the source by Eve (1914), Stahel (1929) and Gray (1936).

§ 8. ESTIMATION OF INTEGRAL DOSE

Suppose we have determined the distribution of dose in röntgens in a medium of unit density and approximately air-like, as, for example, a water phantom. Then the integral dose in a ring element (figure 6) of cross-sectional area da is clearly

$$D \,.\, 2\pi r \,.\, da.$$

The integral dose between the surfaces generated by the rotation of the contours D_1 and D_2 about the central axis is

$$\Sigma = 2\pi \int r \,.\, D \,.\, da. \qquad \ldots\ldots(6)$$

If D is the mean dose between the contours D_1 and D_2, then

$$\Sigma \simeq D \,.\, V, \qquad \ldots\ldots(6\,a)$$

where V is the volume between the isodose surfaces.

To find this volume (Mayneord, 1940a) we suppose the diagram drawn on cardboard, sheet lead or other suitable material having a mass ρ_s per unit area. Then the moment of the shaded area between D_1 and D_2 about the axis AB is

$$\int r da\, \rho_s \,.\, g = -V \,.\, \rho_s g/2\pi. \qquad \ldots\ldots(7)$$

If, therefore, we measure the moment of each section of the contour diagram and multiply by the appropriate conversion factor as well as the mean dose, we obtain an approximation to the total absorption in gm. röntgens between the two surfaces.

To measure the moments we have modified a simple chemical balance (figure 7) by fitting on to it a Perspex platform, the moments being determined by weights in the scale pan. Often comparative measurements are all that are required, but if necessary it is easy, by measurements of the length of the beam of the balance and of ρ_s, to determine the moments and volumes in absolute units.

We have tested the method by measuring the moments of simple shapes which generate on rotation solids whose volumes may be directly calculated.

This method is clearly applicable to all beams of x or gamma rays having an axis of symmetry, and has been used by us (Mayneord, 1940 a) to determine the integral doses delivered by a number of commonly used techniques, and has given results of very considerable significance in relation to the serious constitutional disturbances experienced by patients undergoing high-voltage x-ray or gamma-ray treatment.

There are many other possible methods of estimating integral dose, though very little experimental work has yet been carried out on the subject. For example, we might send our beams into solutions known to be decomposed by x rays and make use of the chemical changes brought about. Grimmett (1939) has constructed a sectional model of a man specially adapted to measure the total energy absorption throughout his mass, and although no results are so far available, this seems a very direct and promising method of attack.

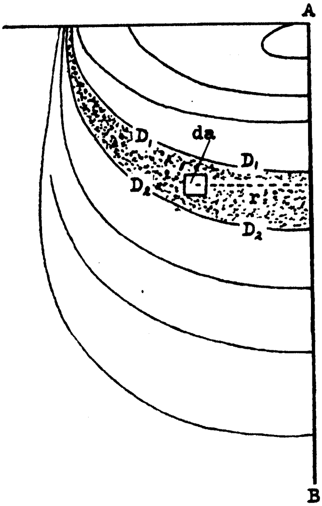


Figure 6.

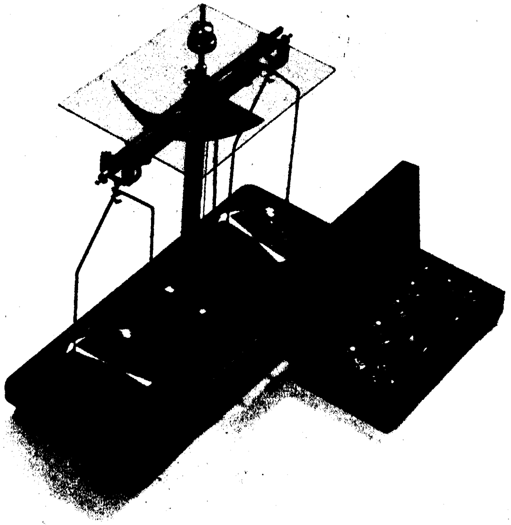


Figure 7.

We have tried to solve the problem in another way by the use of a large number of condenser ionization chambers so placed as to measure the mean dose at the centre of mass of equal masses.* For example, if we suppose the human body to be equivalent in cross-section to an ellipse of semi-axes 18 cm. and 10 cm. respectively, we may divide this ellipse into 72 equal areas, as shown in figure 8. Consider a slice of uniform thickness. The areas then represent equal masses. By placing a condenser chamber at the centre of mass of each section we transform our integral into a summation, so that

---

* Condenser chambers suitable for such work, and made by pressing carbon-bakelite mixtures, were demonstrated at the lecture.

$$\int D_m \, . \, dm = \Delta m \, . \sum_{1}^{r} D_m, \qquad \ldots\ldots(8)$$

and it is merely necessary to add together the indications of all the condensers

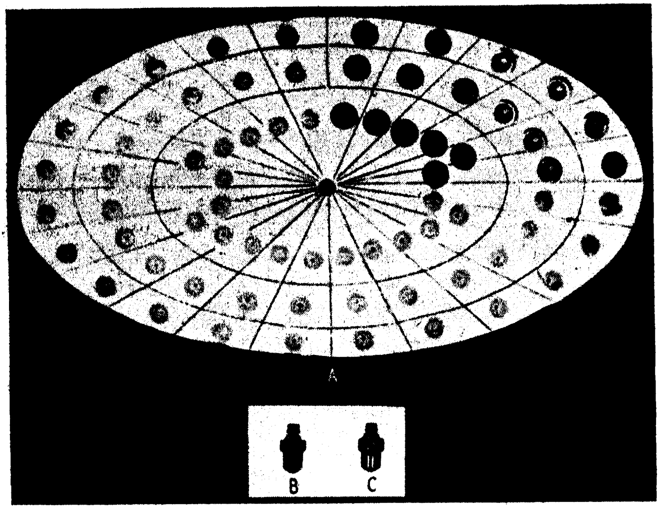


Figure 8. Experimental arrangement for measurement of integral dose.

A. Block of wax divided into 72 equal areas. Ionization chambers in position in one quadrant.
B. Cross-section of condenser chamber.
C. Condenser chamber of pressed bakelite-carbon.

to obtain a quantity proportional to integral dose. By making a number of measurements at different cross-sections we may estimate the total energy absorbed by the body as a whole.

§ 9. SIGNIFICANCE OF INTEGRAL DOSE

It may be wondered why so much interest should be taken in total energy absorption. There are, however, signs that the outlook which has hitherto predominated in dosage problems, namely, that of estimating röntgens at each point within the tissues, has led to an over-simplification of the problem. The energy absorbed by a living cell may not be the sole factor in determining the biological effects of radiation in that cell, for the damage caused to it may also depend on that caused to neighbouring cells which may, in their turn, have been damaged by the radiation. We must remind ourselves that a patient is an organized system, and not a series of disconnected autonomous units.

Moreover, the improvements in precision of the last few years have resulted in raising the dosage levels administered to patients until the severe general constitutional effects of the treatment frequently provide as effective a bar to higher dosage as the local effects previously observed at the skin or other sensitive tissue. The increase of these general effects with the mass of irradiated tissue, that is, size of field and penetration, is a well-known phenomenon of high-voltage x-ray and radium therapy.

Again, when we consider the physical problems involved in the protection of staff and patients from unwanted radiation, the same difficulties are forced upon us. We might guess that the blood changes and general effects of the radiation would depend, not on the dose at a particular point, but on the total energy absorbed by the body as a whole. This is a very important matter needing much more detailed study, for the Protection Regulations lay down the maximum permissible dosage-rate as 10^{-5} r./sec., whereas we have seen that the total energy flowing through a surface per röntgen recorded on it varies very greatly with the wave-length of the radiation. It is chiefly with the object of studying this problem that the condenser-chamber techniques described above have been devised, for it would seem likely that a given surface dose of gamma rays or penetrating x rays would correspond to much greater total energy absorption in the body than the same dose in röntgens of long-wave diagnostic or therapeutic radiations. Again, the obvious dependence of total energy absorbed in the cross-section of the beam merits more interest than has been given to it. Until further measurements have been made, it is difficult to decide on relative total body absorption per unit surface dose, but it should eventually be possible to find the surface doses of different radiations to give roughly constant total energy absorption.

§ 10. APPROXIMATE FORMULAE FOR THE CALCULATION OF INTEGRAL DOSE

It would be a great convenience to have available formulae for the approximate calculation of integral dose in gramme-röntgens when we do not wish to carry out the elaborate detailed analysis, as, for example, in dealing with the integral doses of hundreds of patients as a routine procedure. Formulae of this nature have been given by Happey (1940, 1941) and the writer (1940a, 1940b).

Consider first of all a parallel beam of radiation of cross-section A cm^2 falling normally on the surface of the absorbing medium. We may to a first approximation consider the dose contours rectangular and reaching to the geometrical limits of the beam. Then, if the fall of dose is exponential (as has been found experimentally to be approximately the case in many instances), and D_0 is the dose on the surface, the dose at a depth x cm. will be given by $D_0 e^{-\mu x}$. Then the integral dose Σ to a depth d in a medium of unit density is given by

$$\Sigma = \int_0^d D_0 . A . e^{-\mu x} dx = \frac{A . D_0}{\mu}\left[1 - e^{-\mu d}\right]. \qquad \ldots\ldots(9)$$

If all the energy is absorbed, $\Sigma = D_0 . A . 1/\mu$. ......(10)

It is interesting that $1/\mu$ is the average path of the quanta in the medium. In practice it is more convenient to use a quantity $d_{\frac{1}{2}}$, the depth at which the dose falls to half its surface value. Clearly

$$\mu d_{\frac{1}{2}} = 0{\cdot}693, \qquad \ldots\ldots(11)$$

$$\Sigma = 1{\cdot}44 . D_0 . A . d_{\frac{1}{2}}, \qquad \ldots\ldots(12)$$

a very convenient formula.

If we consider a diverging beam from a point source at a distance f, the absorption in an element at depth x may be assumed to be

$$\Delta\Sigma = D_0 \, . \, A \, . \, e^{-\mu x} \, . \, dx \left(\frac{f+x}{f}\right)^2. \qquad \ldots\ldots(13)$$

Performing the integration, it is found (Mayneord, 1940 a) that for total absorption

$$\Sigma = 1{\cdot}44 \, D_0 \, . \, A \, . \, d_{\frac{1}{2}}\left[1 + \frac{2{\cdot}88 \, d_{\frac{1}{2}}}{f} + \frac{4{\cdot}16 \, d_{\frac{1}{2}}^2}{f}\right]. \qquad \ldots\ldots(14)$$

This is the formula we normally employ in practice in calculations on patients, taking into account the dose on the exit surface if appreciable.

We have shown in the same paper that these formulae give results in fair agreement with the more complex detailed study, but much more precise measurement and calculation are now urgently required.

Clearly the same kind of argument may be used for gamma rays, but we are now more concerned with the possibility of correlating the amounts of energy involved with those used in ultra-violet light and infra-red therapeutic procedures.

§ 11. COMPARISON OF ENERGY FLOW OF X RAYS WITH OTHER RADIATIONS

The measurement of ultra-violet light for medical purposes raises very difficult and complex questions which it would certainly be impossible to discuss here in detail, but it may be said that the difficulty arises largely from the very complex variation of biological effect of the same amount of energy with wave-length. This may be exemplified by the variation of erythema-producing power with wave-length according to the investigations of Coblentz, Stair and Hogue (1932). A sharp maximum, set as 100 %, occurs at $\lambda 2976$, coinciding with a strong line of the hot mercury-arc spectrum. We do, however, have some idea of the absolute amount of energy required to produce an erythema on the skin, for this quantity has been estimated a number of times, with, as might be expected, very varying results. Luckiesh, Halladay and Taylor (1930) give 4300 μ watt sec./cm^2 as the mean of their determinations, while Coblentz, Stair and Hogue (1932) give 20,500 μ watt sec./cm^2 as their mean. Maddock (1936) in a paper to this Society found 11,500 μ watt sec./cm^2 The Council of Physical Therapy of the American Medical Association, based largely on the very extensive pioneer work of Coblentz in these matters, defined a *Finsen Unit* (F.U.) as a flux density of 10 μ watt/cm^2 of homogeneous radiation of wave-length 2976 A., and say that 2 F.U. for 15 minutes is a representative requirement for a minimum perceptible erythema. This would correspond to 18,000 μ watt sec./cm^2 or $1{\cdot}8 \times 10^5$ ergs/cm^2 at the wave-length of maximum sensitivity. If we use the heterogeneous ultra-violet radiation of a mercury arc (high temperature, high vapour pressure, low voltage, quartz tube) shorter than 3130 A., the total energy required is 58 μ watt/cm^2 for 15 minutes, that is about three times as much energy as for the monochromatic radiation (Coblentz, 1940). It is interesting to calculate the doses in röntgens of X rays of different wave-lengths which would be required to give the same energy flux as the minimum perceptible erythema dose in ergs/cm^2 at 2967 A. These values are given below:—

λ in X.U.	Dose in *r* without back-scatter
10	50
50	50
150	54
300	175
800	2,800
1,000	5,700
2,000	40,000

It is of interest that in x-ray techniques in which the whole body is irradiated to produce constitutional effects not unlike those of ultra-violet light, the order of dose normally given is 50 to 100 r. at an average wave-length of 150 to 200 X.U., so that the x-ray energy-flux is of the same order as the energy-flux for ultra-violet. This may, of course, be a mere numerical coincidence, but should be further investigated.

These comparisons are also of interest in relation to the cure of rickets, the production of tumours of the skin by both x rays and ultra-violet, and in many other respects.

§ 12. INFRA-RED RADIATION

It would be expected that the amount of energy administered to a patient during treatment by infra-red radiation would be much larger than in either of the radiation fields hitherto mentioned, but very little investigation of the subject seems to have been made, the dosage of infra-red being in clinical practice almost entirely empirical.

We have so far carried out only preliminary measurements, which may now be described.

A Moll microthermopile with a fluorite window was used in conjunction with a Cambridge galvanometer of short period, the e.m.f. generated by the radiation in the thermopile being compensated in a circuit similar to that used by Guild (1930) for radiation measurements. The thermopile was calibrated at the National Physical Laboratory and found to give 61·5 microvolts per milliwatt of radiant energy. A second calibration is also available, as the Bureau of Standards in Washington very kindly presented us with a standard of thermal radiation in the form of a calibrated carbon-filament lamp of a type supplied by them for such work. A careful comparison of the calibrations is in progress.

We have also available small vacuum thermocouples presented by the General Electric Company (Schenectady) in the form of small glass bulbs containing very thin strip thermo-elements (G.E.C., 1941). These are sensitive, quick acting, and when they were used in conjunction with a low-resistance (70-ohm) microammeter, a current of 51 μa. was obtained for a radiation intensity of 1 g. cal./cm^2 min. These vacuum couples were calibrated by a substitution method, using the Moll thermopile as standard. The instruments were mounted side by side on a short optical bench at right angles to the main optical bench carrying the source of radiation. By moving the instruments along the shorter bench, the General Electric thermoelements could be brought to exactly the position previously occupied by the Moll and so irradiated in similar conditions. The General Electric thermocouple was mounted in a polished metal tube, care being taken to insulate the system thermally so as to reduce conduction to the cold junctions. The thermocouples were calibrated either using the same galvanometer and compensation system or the General Electric instrument connected in series with a low-resistance microammeter, and the currents in this instrument observed directly. The latter procedure is convenient, as the mounted thermocouple and microammeter form a useful clinical instrument with a scale which is very nearly linear up to 100 μa., though there is some departure at higher intensities. Experiments on the wave-length dependence of the calibration are in progress.

Measurements have been made on a number of patients and in conditions likely to occur in practice, the source of radiation during calibration and subsequent use being the same. The results show the urgent need for standardization, since apparently small changes in the source cause large changes in the intensities, while measurements over the surface of the patient or along the axes of the beams reveal very complex variations.

To quote representative results, we have found that, using a 230-volt 500-watt tungsten-filament lamp in its reflector, the amount of radiation received on the skin at a metre distance may vary from 0·58 to 1·8 gm. cal./cm.^{2} min. The normal value employed for such a source is about 1·5 gm. cal./cm.^{2} min. We have also found that an intensity of the order of 2 gm. cal./cm.^{2} min. is about the maximum for comfort, while the threshold of pain is found at about 3·5 gm. cal./cm.^{2} min., in agreement with Hartley (1934), who gave a value of 3·7 gm. cal./cm.^{2} min. for a similar source. The value depends, of course, on the part of the body irradiated. The amount which may be tolerated is, however, very variable with the colour temperature of the source, as has been demonstrated by Sonne (1926), who found that the maximum amounts of visible, "inner" and "outer" infra-red which could be borne by human skin were in the ratios of 3·11 to 1·79 to 1·33 gm. cal./cm.^{2} min., though taking into account reflection, the amounts absorbed were found to be 2·02, 1·16 and 1·33 gm. cal. respectively. These variations may be correlated with the differences in penetration of the radiations. It is interesting to recall that the intensity of sunlight in a clear day in summer at noon is of the order of 1·0 to 1·5 g. cal./cm.^{2} min.

We hope, in conjunction with our medical colleagues, to continue these measurements on a large number of patients, and so obtain information on the many questions which now arise.

Finally, it may be of interest to see the order of amounts of energy used in the various methods.

(a) Type of radiation		Intensity in ergs/cm.^{2} sec.
High voltage x rays. 200 kv. Heavy filter	40 r./min.	2×10^3
Ultra-violet light	2 Finsen units 20 μ watt/cm.^{2}	2×10^2
Infra-red	1 gm. cal./cm.^{2} min.	7×10^5

(*b*) Type of radiation		Total energy absorbed.
x rays. 200 kv.	30 megagramme-röntgens	60 gm. cal. in six weeks
Ultra-violet light	Erythema over body surface of 1500 cm^2	8 gm. cal. at one sitting
Infra-red	1 gm. cal./cm^2 min. 20 min. Area 1500 cm^2	3×10^4 gm. cal. at one sitting

§ 13. CONCLUSIONS

We have seen how the measurement of radiation for medical purposes is concerned primarily with two energy-absorption problems, first the energy absorbed at each point in the tissues, and secondly the energy absorbed in the body as a whole. Concerning the first problem, in the x-ray and gamma-ray field we now know a little, but so far as ultra-violet and infra-red are concerned, practically nothing.

As regards the total energy absorbed by the body as a whole, we have now rough estimates which are of great interest and enable us to link up many different problems.

Only on an energy basis can these problems be correlated, and it is hoped that this sketch of a very wide field may do something to stimulate interest in the measurement of radiation in absolute units for medical purposes.

REFERENCES

Bragg, W. H., 1912. *Studies in Radioactivity* (Macmillan and Co., Ltd.).
Coblentz, W. W., 1940. *J. Amer. Med. Assoc.* **114**, 325.
Coblentz, W. W., Stair, R. and Hogue, J. M., 1932. *Bur. Stand. J. Res., Wash.*, **8**, 541, 759.
Eisl, A., 1929. *Ann. Phys., Lpz.*, **3**, 227.
Eve, A. S., 1914. *Phil. Mag.* **27**, 394.
G.E.C., 1941. *Gen. Elect. Rev.* **44**, 133.
Gray, L. H., 1936. *Proc. Roy. Soc.* A, **156**, 578.
Gray, L. H., 1937. *Brit. J. Radiol.* **10**, 600; **10**, 721.
Grimmett, L. G., 1939. *Amer. J. Roent.* **41**, 432.
Guild, J., 1930. *J. Sci. Instrum.* **7**, 378.
Happey, F., 1940. *Nature, Lond.*, **145**, 668; **146**, 96; 1941. *Brit. J. Radiol.* **14**, 235.
Hartley, H., 1934. *J. Inst. Fuel*, April.
Honeyburne, J., Lamerton, L. F., Smithers, D. W. and Mayneord, W. V., 1939. *Brit J. Radiol.* **12**, 262.
Kaye, G. W. C., Bell, G. E., Binks, W. and Perry, W. E., 1939. *Rep. Progr. Phys.* **6**, 95.
Lamerton, L. F., 1940. *Brit. J. Radiol.* **13**, 136.
Lauritsen, C. C., 1933. *Amer. J. Roent.* **30**, 380, 529.
Luckiesh, M., Halladay, L. L. and Taylor, A. H., 1930. *J. Opt. Soc. Amer.* **20**, 423.
Maddock, A. J., 1936. *Proc. Phys. Soc.* **48**, 57.
Mayneord, W. V., 1933. *Brit. J. Radiol.* **6**, 598.
Mayneord, W. V., 1938. *Rep. Progr. Phys.* **5**, 284.
Mayneord, W. V., 1940 a. *Brit. J. Radiol.* **13**, 235; 1940 b. *Nature, Lond.*, **145**, 972.
Mayneord, W. V. and Honeyburne, J., 1941. *Amer. J. Roent.* **45**, 235.
Mayneord, W. V. and Lamerton, L. F., 1941. *Brit. J. Radiol.* **14**, 255.
Mayneord, W. V. and Roberts, J. E., 1937. *Brit. J. Radiol.* **10**, 365.
Sonne, C., 1926. *Acta radiol., Stockh.*, **5**, 419.
Stahel, E., 1929. *Strahlentherapie*, **33**, 296.
Taylor, L. S. and Singer, G., 1940. *Amer. J. Roent.* **44**, 428.

DIELECTRIC LOSS IN PARAFFIN-WAX SOLUTIONS *

By H. FRÖHLICH,
H. H. Wills Physical Laboratory, University of Bristol

MS. received 7 *March* 1942

ABSTRACT. In this paper, the dielectric loss of a dilute solution of long-chain molecules, containing dipoles, in solid paraffin is investigated theoretically. In such substances, realized experimentally by Jackson, Sillars and Pelmore, the interaction between dipoles is negligible, and they thus follow the Debye theory. In order to account for the time of relaxation τ, the structure of these substances is investigated and it is shown that only the case where the dipole-containing chain molecule is shorter than the paraffin molecule can be expected to yield simple results. In order to obtain τ, the minimum energy V required to turn a molecule from one equilibrium position to another has been calculated. For short molecules, V is mainly determined by the potential hill between two equilibrium positions, whereas for long molecules an internal torsion of the molecule becomes important. V is found to increase linearly with chain length for very short molecules, and to tend towards an asymptotic value for long chains. Good agreement with experiments, so far as available at present, has been reached.

§ 1. INTRODUCTION

It is well known experimentally that dielectric losses are very small in solid insulators which form ionic or atomic crystals. Appreciable dielectric losses occur mainly in molecular crystals, built up of molecules containing electrical dipoles.

The fact that dipoles are in many cases responsible for dielectric losses was recognized by Debye (1929) and led to his well-known theory of this effect. Applied to solids, the main idea of this theory is as follows: suppose that each dipole has two positions of equilibrium corresponding to opposite dipole directions with equal energy, and a certain transition probability from one to the other. Now assume that an external electric field is being applied. Then, clearly, some of the dipoles will turn into more favourable directions, thus giving rise to a dielectric polarization. Owing to the finite transition probability, it will take some time to reach equilibrium. This time is measured by the time of relaxation τ. Obviously, if the external field is periodic, the polarization will also be periodic, but with a certain shift of phase. This, according to Maxwell's equations, gives rise to dielectric losses.

Although there seems to be little doubt that the general ideas of Debye's theory are correct, detailed comparisons with the results of experiments on solids have met with varying success. This is probably due to the fact that the interaction between dipoles has not been described in a satisfactory way.

---

* Based on Report L/T. 121 of the British Electrical and Allied Industries Research Association (E.R.A.).

In order to measure the time of relaxation τ of chain molecules without meeting the difficulties just mentioned, W. Jackson (1935) diluted polar long-chain molecules in paraffin wax. Pure paraffin wax, being non-polar, shows practically no dielectric loss. The losses in the solid solution, therefore, can be ascribed to the dissolved polar molecules. Moreover, if their concentration is small enough, their interaction will be negligible.

In this way, Jackson (1935), and later R. W. Sillars (1938) and D. R. Pelmore (1939), determined the time of relaxation of long-chain esters of various chain lengths. Their results show that with increasing chain length, $\log \tau$ increases too, but slower than linearly, and probably approaches an upper limit for very long chains. Now the significance of $\log \tau$ is as follows: Let V_n be the minimum energy which is required in order to lift a molecule with n links over the potential hill between the two stable positions. Then, since $1/\tau$ represents the transition probability, it may be expected that

$$\tau = Ce^{V_n/kT}, \qquad \ldots\ldots(1)$$

where, compared with $e^{V_n/kT}$, C varies only slowly with temperature and chain length. Assuming the molecule to be rigid, V_n, and hence $\log \tau$, must be proportional to the chain length. The experiments, therefore, show that this is not the case, but that for long chains V_n approaches a constant value, independent of the chain length.

It is the aim of this paper to calculate the dependence of V_n on the chain length. Allowing for internal rotational deformations of the molecule, it will be shown that this dependence has the simple form of a hyperbolic tangent. The result, which is demonstrated in figure 3, shows good agreement with experiment. It ought to be mentioned that this result holds only if the dipolar molecule is shorter than the paraffin molecule. In § 2 this will be derived from a study of the lattice structure of paraffins. In § 3, V_n will be calculated, and the results will be discussed in § 4.

§ 2. STRUCTURE

We shall start with a discussion of the possible positions which the long-chain ester molecules may be expected to take under the assumption that they form a real solid solution in the paraffin crystal. The structure of paraffins is well known from the work of A. Müller (1928, 1932). To describe it, we divide the crystal into plane layers, each of the thickness of the chain length. Within a layer, the unit cell is rectangular, with sides of lengths b_1, b_2 and c. b_1, which is of the order of 7·5 A., and b_2 (5 A.) are parallel to the surface of the layer ; c is perpendicular to it, and its length is equal to that of the chain. All chains are parallel to the c-axis. Figure 1 shows the way in which they cross the b_1b_2-axis. It is very important to notice that, in the subsequent layer, the whole arrangement is shifted by about 1 A. in the b_1 direction. These atoms are indicated by dotted lines in figure 1. Now let us try to build a foreign chain into the paraffin crystal. This chain must obviously be parallel to the paraffin chains. It may either replace a paraffin chain, or it may be in an interstitial position, probably in the middle of the b_1-axis. It does not seem possible to decide at present which of the two alternatives is realized, but for the following considerations this is unimportant. Now clearly we have to distinguish two cases, one where the length of the foreign molecule is shorter and the other where it is longer than the paraffin chain. In the first case, the foreign molecule can be built in between the paraffin chains of one layer without being deformed. In the second case this is no longer possible, because of the above-mentioned displacement of two successive layers relative to each other. Such a long molecule has, therefore, to be deformed, in order to be built into the crystal. There is some experimental evidence for the difference between the two cases. X-ray measurements by Piper and others (1931, 1937) show that the crystal spacing is hardly changed if we add to a crystal of chain molecules a small concentration of shorter molecules, whereas, in contrast, the addition of longer molecules increases the spacing considerably.

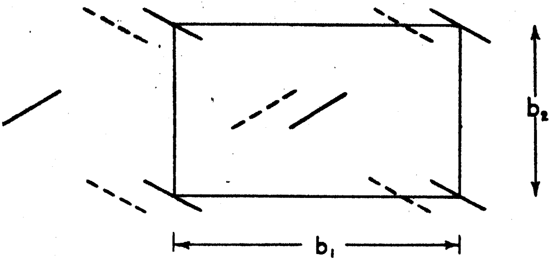


Figure 1. Structure of paraffins perpendicular to chain axis.

Now let us consider the equilibrium position of a chain in more detail. Any such chain forms a plane zigzag, and, as indicated in figure 1, these zigzag planes are in certain fixed directions.* The same must hold for the zigzag plane of a foreign molecule. If such a molecule contains a dipole, then there are at least two equilibrium positions with equal energy but different dipole directions. This is shown in figure 2 under the assumption that the foreign molecule replaces a paraffin chain. The arrows indicate the projection of the dipole direction on the b_1b_2 plane.

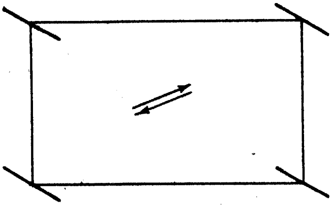


Figure 2. The two equilibrium positions of a dipolar chain.

§ 3. CALCULATION OF V_n

Let us try to turn the dipolar molecule from one equilibrium position to the other, by rotating it at one end. Then, assuming it to be rigid, we have at the same time to lift the whole molecule to the top of the potential hill which separates the two stable positions. If, however, we allow for an internal elasticity, the molecule will be twisted. In order to simplify the following calculations, we shall treat the chain as a continuum, described by elastic constants.

---

* At high temperatures, slightly below the melting point, many paraffins change into a modification in which the chains probably rotate. We shall restrict ourselves, however, to the more important low-temperature modifications.

Let z be the distance of any point of the chain axis from one end, and suppose that a CH_2 group with the co-ordinate z has been turned from one of its equilibrium positions by an angle ϕ around the z-axis. Assuming the chain to be a continuum, ϕ is a function $\phi(z)$ of z, and we shall have to calculate the energy of a deformation in terms of ϕ. Thus, the potential energy will be $\propto \phi^2$, owing to the interaction with its neighbouring molecules. If ϕ depends on z, the chain is twisted, and there will be an additional internal elastic energy $\propto \left(\frac{d\phi}{dz}\right)^2$. Thus the total energy is given by

$$\int_0^{l_n} \left\{ \frac{A^2a^2}{2}\left(\frac{d\phi}{dz}\right)^2 + \frac{B^2}{2}\phi^2 \right\} \frac{dz}{a}. \qquad \ldots\ldots(2)$$

Here l_n is the chain length, and a is the distance between two links, i.e.

$$an = l_n; \qquad \ldots\ldots(3)$$

A^2 and B^2 have the dimensions of an energy, whose order of magnitude we shall estimate below. According to figure 2, with $\phi = 0$, $\phi = \pi$ is also an equilibrium position of the chain. The energy of a small deformation from this latter position is obtained from (2) by replacing ϕ by $\phi - \pi$.

In order to be able to carry out quantitative calculations, we shall assume the ϕ^2 law to hold in the interval $-\frac{1}{2}\pi \leqslant \phi \leqslant \frac{1}{2}\pi$ and the $(\phi-\pi)^2$ law in $\frac{1}{2}\pi \leqslant \phi \leqslant \frac{1}{2}\pi$. Then the interaction energy has its maximum at $\pi/2$ and at $-\pi/2(=3\pi/2)$. Now the most general position of the molecule having an energy near V_n must be such that $\phi = \pi/2$ at a certain $z = z_0$ with both ends $z = 0$ and $z = l_n$ hanging down towards $\phi = 0$ and $\phi = \pi$. The potential energy W of the molecule in such a position is given by

$$W = \int_0^{z_0} \left\{ \frac{A^2a^2}{2}\left(\frac{d\phi}{dz}\right)^2 + \frac{B^2}{2}\phi^2 \right\} \frac{dz}{a} + \int_{z_0}^{l_n} \left\{ \frac{A^2a^2}{2}\left(\frac{d\phi}{dz}\right)^2 + \frac{B^2}{2}(\phi-\pi)^2 \right\} \frac{dz}{a}. \qquad \ldots\ldots(4)$$

V_n is to be determined from the condition that W is stationary, i.e.,

$$\delta W = 0. \qquad \ldots\ldots(5)$$

The variation of W comprises a variation of the function $\phi(z)$ as well as of z_0. ϕ has to be determined so as to make W a minimum for any given z_0. Then, starting to turn the chain at $z_0 = l_n$, i.e. lifting the one end of the chain from $\phi = 0$ up to $\phi = \pi/2$, we shall move z_0 towards $z = 0$, which means that we gradually move the chain over the potential hill, always using the most suitable $\phi(z)$. It follows at once from symmetry that the energy thus determined reaches its maximum at $z_0 = l_n/2$. Hence V_n is obtained from (5) with this value of z_0. Using this value of z_0, we obtain from (5) and (4)

$$a^2A^2\frac{d^2\phi}{dz^2} - B^2\phi = 0, \text{ if } 0 \leqslant z \leqslant \frac{l_n}{2}, \qquad \ldots\ldots(6a)$$

and

$$a^2A^2\frac{d^2\phi}{dz^2} - B^2(\phi-\pi) = 0, \text{ if } \frac{l_n}{2} \leqslant z \leqslant l_n, \qquad \ldots\ldots(6b)$$

with the condition that

$$\phi=\pi/2 \text{ at } z=l_n/2, \qquad \ldots\ldots(7)$$

and the condition for free ends, viz.,

$$\frac{d\phi}{dz}=0 \text{ at } z=0, \text{ and } z=l_n. \qquad \ldots\ldots(8)$$

We can now calculate V_n, which is given by

$$V_n=\int_0^{l_n/2}\left\{\frac{A^2a^2}{2}\left(\frac{d\phi}{dz}\right)^2+\frac{B^2}{2}\phi^2\right\}\frac{dz}{a}+\int_{l_n/2}^{l_n}\left\{\frac{A^2a^2}{2}\left(\frac{d\phi}{dz}\right)^2+\frac{B^2}{2}\phi^2\right\}\frac{dz}{a},$$

where ϕ satisfies (6), (7) and (8). Using these equations, we find by a partial integration

$$V_n=A^2a\left(\phi\frac{d\phi}{dz}\right)_{z=l_n/2}=A^2a\frac{\pi}{2}\left(\frac{d\phi}{dz}\right)_{z=l_n/2}. \qquad \ldots\ldots(9)$$

To determine $d\phi/dz$ we have to solve the differential equations (6). Using the boundary conditions (7) and (8) we obtain

$$\phi=\frac{\pi}{2}\frac{e^{z/an}+e^{-z/a\bar{n}}}{e^{n/2\bar{n}}+e^{-n/2\bar{n}}}, \text{ if } 0\leqslant z\leqslant l_n/2,$$

and

$$\pi-\phi=\frac{\pi}{2}\frac{e^{(z-an)/a\bar{n}}+e^{-(z-an)/a\bar{n}}}{e^{n/2\bar{n}}+e^{-n/2\bar{n}}}, \text{ if } \frac{l_n}{2}\leqslant z\leqslant l_n,$$

where

$$\bar{n}=A/B. \qquad \ldots\ldots(10)$$

Therefore

$$\left(\frac{d\phi}{dz}\right)_{z=l_n/2}=\frac{\pi}{2a\bar{n}}\frac{e^{n/2\bar{n}}-e^{-n/2\bar{n}}}{e^{n/2\bar{n}}+e^{-n/2\bar{n}}}.$$

Inserting this value into (9), we finally obtain

$$V_n=\frac{\pi^2}{4}AB\tanh\frac{n}{2\bar{n}}=\frac{\pi^2}{4}B^2\bar{n}\tanh\frac{n}{2\bar{n}}. \qquad \ldots\ldots(11)$$

We thus see that, for short chains, $n\ll\bar{n}$, V_n becomes

$$V_n\simeq\frac{B^2n}{2}\frac{\pi^2}{4}, \quad n\ll n,$$

as would be obtained if the molecule were considered to be rigid. For long chains, $n\gg\bar{n}$, however, V_n becomes

$$V_n\simeq AB\,\pi^2/4, \quad n\gg\bar{n},$$

independent of the chain length. In this case, therefore, the torsion of the chain is of importance.

§ 4. DISCUSSION

If we insert equation (11) into (1) we find for the time of relaxation τ

$$\log 2\pi\tau=\alpha+\frac{n^2}{4}\frac{B^2\bar{n}}{kT}\tanh\frac{n}{2\bar{n}}. \qquad \ldots\ldots(12)$$

This formula contains three constants: α, B and $\bar{n}=A/B$. We shall not attempt here to determine α theoretically, but the order of magnitude of B and $\bar{n}$ can be estimated as follows: from its definition, B^2 is the energy required per link in order to turn the untwisted chain by an angle $\phi = 2$ ($\sim\pi/2$). This should be about the same energy as is required per link in order to turn a paraffin chain in the same way. Since at about 300° K. paraffins have a transition into a state in which their molecules can probably rotate, we estimate

$$B^2 \sim 1/30 \text{ ev.} \qquad \ldots\ldots(13)$$

A^2 determines the energy required to twist the chain, i.e. to deform the valency angle. Thus A^2 may be expected to be of the same order of magnitude as the energy of the chemical bond. Thus we expect

$$A^2 \sim 5 \text{ ev.} \qquad \ldots\ldots(14)$$

Hence according to (10)

$$\bar{n} = A/B \sim 10. \qquad \ldots\ldots(15)$$

Before discussing our results in a more quantitative way we must recall that in § 2 we assumed that the dipolar molecule is shorter than the paraffin chain. If it is longer it will have to be strongly deformed in order to be built into the lattice. No simple theoretical significance can be given to such measurements. With these remarks in mind, we shall now compare quantitatively the experimental and theoretical dependence of $\log\tau$ on the chain length.

From Sillars' experiments we know $\log\tau$ for $n = 20$, 22, 24, 28 and 32 in ordinary paraffin wax. Since the chain length of the latter is between 24 and 28, only the values for $n = 20$, 22 and possibly 24 can be used. Measurements by Pelmore (1939) are, however, available for $n = 32$ in very long paraffin chains (up to about 52 links).

In the following table we give the experimental values of $-\log_{10}(2\pi\tau)$ at $-20°$ C. The values in brackets are those which have to be excluded:

n	20	22	24	28	32
$-\log_{10}(2\pi\tau)$	$+6{\cdot}7$	$+5{\cdot}5$	$4{\cdot}8$	$(4{\cdot}2)$	(4) (Sillars) 2·3 (Pelmore)

In order to determine the unknown constants in the theoretical formula (12) we shall use the experiments on the temperature dependence of τ. Unfortunately, for the substances for which our theory holds, the measured temperature range is too small to check the theoretical temperature dependence. Assuming it to be correct, it is possible to determine (from $n = 20$) two constants, namely,

$$\alpha = -50{\cdot}4, \quad \frac{B^2\bar{n}}{k} = 5200°. \qquad \ldots\ldots(16)$$

The remaining constant $\bar{n}$ can be determined from the ratio of τ for $n = 20$ and for $n = 22$, which gives

$$\bar{n} = 13. \qquad \ldots\ldots(17)$$

Thus from (16) and (10) the values of A^2 and B^2 are

$$A^2 \sim 6 \text{ ev.}, \quad B^2 \sim 1/30 \text{ ev.}, \qquad \ldots\ldots(18)$$

in agreement with the theoretically estimated order of magnitude (14) and (13).

In figure 3 we have plotted the theoretical curve, i.e. $\tanh n/2\bar{n}$, as a function of $n/2\bar{n}$. Using the above values for α, A^2 and B^2, we see that the experimental points for $n=24$ and $n=32$ lie on the theoretical curve, whereas the values for $n = 28$ and 32, which had to be excluded (in brackets), fall below.

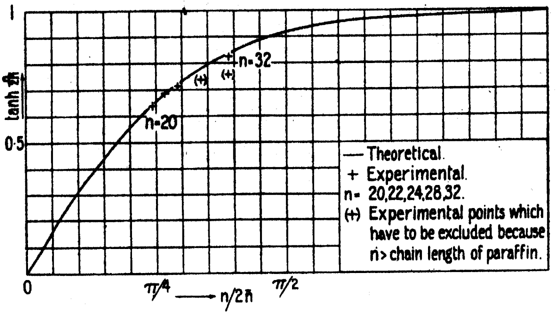


Figure 3. Dependence of log τ on chain length.

The agreement reached can be considered as a confirmation of our formula (12). It is satisfactory to notice that the value of $B^2\bar{n}$ obtained from the temperature dependence agrees with that obtained from the dependence on chain length, and has, moreover, the theoretically estimated order of magnitude.

REFERENCES

Debye, 1929. *Polare Molekeln* (Leipzig).
Francis, Collins and Piper, 1937. *Proc. Roy. Soc.* A, **158**, 691.
Jackson, W., 1935. *Proc. Roy. Soc.* A, **150**, 197.
Müller, A., 1928. *Proc. Roy. Soc.* A, **120**, 437 ; 1932. *Ibid.* **130**, 514.
Pelmore, D. R., 1939. *Proc. Roy. Soc.* A, **172**, 502.
Piper, Chibnall, Hopkins, Smith and Williams, 1931. *Biochem. J.* **25**, 2072.
Sillars, R. W., 1938. *Proc. Roy. Soc.* A, **169**, 66.

THE DIMENSIONS OF PHYSICAL QUANTITIES

By E. W. H. SELWYN

Communication H849 from the Kodak Research Laboratories, Harrow

Received 20 *March* 1942

ABSTRACT. It is shown that the simple form of relation between physical quantities which is assumed in considerations of the dimensions of physical quantities may be regarded as a consequence of the equality of two measures of a physical quantity, on any system of units, if they are equal on one system. The argument is extended to cover the "method of dimensions".

§ 1. INTRODUCTION

A few years ago, in a theoretical discussion of the origin of the graininess of photographic materials (Selwyn, 1935), an argument was developed which involved the supposition that photographic density could be regarded as having a dimension in a way similar to that in which mass or time has a dimension. Now photographic density is defined as the logarithm of the ratio of two light fluxes or even of two transmissions. Thus, in the ordinary way, it would undoubtedly be regarded as of zero dimensions. But the assumption that it had a dimension (namely, the dimension of density) did, in fact, lead to an answer in agreement with answers obtained by straightforward physical arguments, and, indeed, in substantial agreement with observation. Moreover, the form of the expression arrived at, by whatever process, forces upon one the conclusion that the " graininess constant " which forms a convenient measure of the granularity of photographic materials is of the dimensions [*photographic density*] . [*length*]. It was felt, by analogy with certain other physical " constants ", such as specific heat, or the quantity a in the equation

$$T = 2\pi\sqrt{l/g}(1 + a\theta^2)$$

for the period of a pendulum vibrating through a small finite arc θ, which has different values accordingly as the amplitude is measured in radians or degrees, that there was justification enough for this assumption that density had a dimension of its own. The problem was therefore worked out by the ordinary " method of dimensions ". Then Bridgman's *Dimensional Analysis* (Bridgman, 1931) came to the author's notice, and the arguments there put forward were immediately adapted to suit the special case in question, since they appeared much more convincing. The treatment which now follows of the general question of dimensions is broadly that followed by Bridgman,* but it is believed to have sufficient points of novelty, particularly in the derivation of the fundamental formulae, to justify publication.

§ 2. THE FUNDAMENTAL FORMULAE

In measuring a physical quantity, the process adopted is to measure certain other quantities, say, a, b, c, d, ... , and combine them by a specified method of calculation to yield the desired result, say p. Thus

$$p = F\ (a, b, c, d, \ldots.). \qquad \ldots\ldots(1)$$

If the quantity being measured is a simple one, such as a length for instance, the process of measurement and calculation is very simple. One may regard the process of measuring a length as consisting in taking a rod and finding out how many times the rod can be placed end to end within the length which is being measured. The number of times so counted is called the length, while the rod with which the measurement is made is called the unit of length. Obviously, if the unit were shortened to a fraction $1/L$ of its original length, the number which represents the measured length would be multiplied by L. If we suppose that all the quantities in the brackets of (1) are such that this reciprocal relation exists between the measures and the units, and we change the units in the ratios $1/A$ for the a-quantity, $1/B$ for the b-quantity, and so on, we shall find a new value for the p-quantity, namely,

$$p' = F(Aa, Bb, Cc, Dd \ldots.). \qquad \ldots\ldots(2)$$

It is customary to repeat physical measurements several times, and we now suppose that a second measurement of p has been made, with different values a', b', c', d', ... as the starting point, but with the result remaining the same as before, namely p. We now impose the very reasonable condition that the result of the calculation after the units of the a'-, b'-, c'-, d'-, ... quantities have been changed in the ratios $1/A$, $1/B$, ... shall remain p'. That is, we make the requirement that if two measurements of a physical quantity are found to be equal on one system of units, they shall remain equal on any other system. It would be difficult to find a less restrictive condition. In other words, we expect that

---

* Resemblances will also be found to that given by Campbell (1920).

$$\left.\begin{aligned} &\text{if} && F(a, b, c, d, \ldots.) \text{ is constant,} \\ &\text{so also is} && F(Aa, Bb, Cc, Dd, \ldots.), \end{aligned}\right\} \qquad \ldots\ldots(3)$$

whatever the values of A, B, C, D. This constitutes a limitation and, as will be seen immediately, a very considerable limitation on the form of F, that is upon the types of calculation permissible in deriving physical quantities.

Let us consider the simplest case, in which

$$p = F(a, b) \qquad \ldots\ldots(4)$$

$$\text{and} \quad p' = F(Aa, Bb). \qquad \ldots\ldots(5)$$

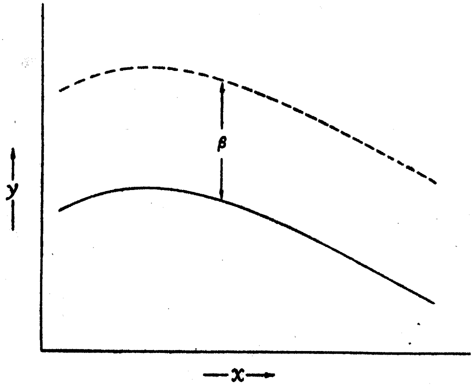
First change the variables to $x = \log a$, $y = \log b$, $\alpha = \log A$, $\beta = \log B$ and $p = z, p' = z'$. Then we can write the above equations as

$$z = f(x, y),$$
$$z' = f(x + \alpha, y + \beta).$$

We consider these expressions to represent surfaces whose height above the plane of the x-y axes is z, and in the figure we take the full line to represent the contour line

$$f(x, y) = z_0.$$

Now for every point (x, y) on this curve, draw a corresponding point $(x, y + \beta)$. The curve through the points $(x, y + \beta)$ will obviously be parallel to the curve through the points (x, y). But according to the preceding arguments, if $f(x, y)$ is constant, so also is $f(x, y + \beta)$. That is, since $f(x, y)$ or $f(x, y + \beta)$ is the height of the surface above the plane of the x-y axes, the second curve is also a contour line. Thus, by these arguments, either z is independent of x and y (a trivial solution), or all contour lines are parallel. The same arguments can be applied



to changes of *a*-units, and also to simultaneous changes of both *a*- and *b*-units. These conditions are clearly only satisfied if the contour lines are straight, for with any parallel lines which are not straight, a displacement can be found which will cause them to intersect. Thus

$$z = f(lx + my), \qquad \ldots\ldots(6)$$

where l and m are constants, and f represents an arbitrary function.

An exactly similar argument can be used in discussing the "contour-surfaces" involved in the four-dimensional case, for which the result is

$$z = f(lx + my + nz). \qquad \ldots\ldots(7)$$

This result can be extended to any number of co-ordinates by taking them in groups of any three, while maintaining the remainder constant. Since all groups of any three must follow the preceding equation, the final result after changing back to the original variables is that

$$p = F(a^{\alpha}\, b^{\beta}\, c^{\gamma}\, d^{\delta}\, \ldots.). \qquad \ldots\ldots(8)$$

Bridgman suggests that it is required of physical quantities that they shall preserve the same ratio when the units of measurement are changed. We can arrange for this by taking the inverse function, F^{-1}, of p as the measure of the quantity in question, for then its measure becomes

$$a^{\alpha}\, b^{\beta}\, c^{\gamma}\, d^{\delta}\, \ldots. \qquad \ldots\ldots(9)$$

and in different units

$$A^{\alpha}\, B^{\beta}\, C^{\gamma}\, D^{\delta}\, \ldots.\ a^{\alpha}\, b^{\beta}\, c^{\gamma}\, d^{\delta}\, \ldots. \qquad \ldots\ldots(10)$$

The convenience of this convention is obvious. There seems to be no very fundamental reason, however, why we should not find it convenient, on occasion, to use quantities defined by

$$F(a^{\alpha}\, b^{\beta}\, c^{\gamma}\, d^{\delta}\, \ldots.).$$

In fact, it is common to do so. For some purposes, for instance, the logarithm of a brightness is a better unit than brightness itself. But it must be remembered that if brightness measurements are to be combined with other measurements in deriving a physical quantity, it will be found simpler to use brightness itself rather than its logarithm. The *a*-, *b*-, *c*-, *d*-quantities involved in the derivation of a *p*-quantity are therefore themselves invariably reduced to such a form that they can be expressed by formulae of the type of (9) before they are used in the calculation. Moreover, so long as the form of F is known, it is immaterial whether we calculate the quantity we are interested in as $a^{\alpha}\, b^{\beta}\, c^{\gamma}\, d^{\delta}\, \ldots.$ or as $F(a^{\alpha}\, b^{\beta}\, c\;\, d^{\delta}\ldots.)$. From this point, therefore, we shall assume the simpler form.

The above demonstration implicitly assumes that a, b, c, d, ... are all independent. But it may happen that the units of the individual components of a group of quantities, say $e^{\epsilon} f^{\zeta} g^{\eta}$, are such that $E^{\epsilon} F^{\zeta} G^{\eta}$ is independent of the values of E, F, G. This occurs when e, f, g have component measures in common, the indices of which in the product $e^{\epsilon} f^{\zeta} g^{\eta}$ all add to zero. A product of this sort is termed dimensionless, because a change in the size of the units involved in it is without effect on its value. Consequently, as a little consideration will show, in these circumstances,

$$p = a^{\alpha} b^{\beta} c^{\gamma} d^{\delta} \dots \phi (e^{\epsilon} f^{\zeta} g^{\eta}),$$

and in general

$$p = a^{\alpha} b^{\beta} c^{\gamma} d^{\delta} \dots \phi (a', b', c', d' \dots),$$

where the arguments a', b', c', d', ... of ϕ are each dimensionless products similar to $e^{\epsilon} f^{\zeta} g^{\eta}$. These products must, of course, be independent.

Some p-quantities are constant. Typical examples are

$$G = d^2 F / m_1 m_2,$$

where d is the distance between two masses m_1, m_2 and F the force of attraction between them; and

$$a = u / \Theta^4,$$

where u is the energy density of radiation in a black-body enclosure at absolute temperature Θ. Constants such as these have different values according to the units used in the measurement of the component factors, and are commonly known as dimensional constants. Other constants, such as

$$\pi = C/D,$$

where C is the circumference and D the diameter of circular objects, and "Reynolds number",

$$v \rho d / \eta,$$

where v is the velocity of a fluid, η the viscosity, ρ the density of the fluid, and d a linear dimension describing the size of the channel in which the fluid is flowing, are independent of the units used in its measurement, and may be conveniently termed "function constants". Now these constants may be regarded as substitutes for certain measurements and must, therefore, be included, in certain fairly obvious circumstances, in the formula for p. To take a very simple example, we might be interested to know the relation between the diameter of a wheel and the distance travelled by it in so many turns. The formula for the calculation involves π, because the quantity we measure is D, but the quantity which determines the distance travelled is C. In the same way, if we wish to know the volume of a liquid drop which falls from a tube, among other things we know that one of the important factors is the force on the drop tending to pull it out of the tube. We may measure this in our experiments with the drop itself by finding how fast it falls and the weight of the drop, or by making use of the constant g for gravitational acceleration. There is, however, a fundamental difference between the two sorts of constant. A "function constant" is necessarily included among the arguments of ϕ, but a dimensional constant can be involved in ϕ only if it is combined with such other terms as to make the resultant dimensionless.

We have now reached the general result that the complete formula for a physical quantity is

$$A^{\alpha} B^{\beta} C^{\gamma} D^{\delta} \dots a^{\alpha} b^{\beta} c^{\gamma} d^{\delta} \dots \phi (a', b', c', d' \dots),$$

wherein $A^{\alpha} B^{\beta} C^{\gamma} D^{\delta}$ .... represents what are known as the "dimensions" of the quantity, but where the rest of the formula represents what "physical reality " the quantity has. The insignificance of $A^{\alpha} B^{\beta} C^{\gamma} D^{\delta} \ldots .$ is emphasized by the fact that it is hardly ever left in physical equations. There are two reasons why it is left out. One is that there are no absolute units by which $a, b, c, d \ldots .$ might be measured. One set of units is therefore just as appropriately used for $a, b, c, d \ldots .$ as another. The other is that it can always be determined without the slightest difficulty from the normal formula

$$p = a^{\alpha} b^{\beta} c^{\gamma} d^{\delta} \ldots . \phi (a', b', c', d' \ldots .).$$

§ 3. THE METHOD OF DIMENSIONS

A reverse operation is possible to some extent. This constitutes the basis of the " method of dimensions ". The fundamental assumption underlying this process of arriving at relations between physical quantities is that the dimensional formula for the p-quantity may be known independently of the way in which the relevant factors determine it in the experiment actually under consideration. This is, indeed, a necessary condition for most of the operations of physics. For the experimental determination of the relation between p and its component measurements $a, b, c, \ldots .$ presupposes a knowledge of p by other means. We might, for instance, define the surface tension of a liquid by means of measurements of the volume of drops falling from a tube, but this would mean nothing very much unless the same results (within reasonable limits) were also secured for surface tension by measurements on ripples on the surface of the liquid or the height of ascent of the liquid in capillary tubes. Thus we may start off with a knowledge of the dimensions of p in terms of certain units a_0, b_0, c_0, d_0, and the suspicion that p is also measurable in terms of the units $a, b, c, d, \ldots .$ We know that if this is so

$$p = a^{\alpha} b^{\beta} c^{\gamma} \ldots . \phi (a, b, c, \ldots .),$$

and we also know the measures of $a, b, c \ldots .$ in terms of $a_0, b_0, c_0 \ldots .$, namely,

$$\begin{aligned} a &= a_0^{\alpha_1} b_0^{\beta_1} c_0^{\gamma_1} \ldots . \\ b &= a_0^{\alpha_2} b_0^{\beta_2} c_0^{\gamma_2} \ldots . \\ &\text{etc.} \end{aligned}$$

Thus,

$$p = (a_0^{\alpha_1} b_0^{\beta_1} c_0^{\gamma_1} \ldots .)^{\alpha} (a_0^{\alpha} b_0^{\beta_2} c_0^{\gamma_2} \ldots .)^{\beta} \ldots . \phi (a, b, c, \ldots .),$$

and the dimensional part of the formula for p is

$$(A_0^{\alpha_1} B_0^{\beta_1} C_0^{\gamma_1} \ldots .)^{\alpha} (A_0^{\alpha_2} B_0^{\beta_2} C_0^{\gamma_2} \ldots .)^{\beta} \ldots .$$

The process of the " method of dimensions " is then to make this formula identical with the known dimensional formula for p. This is done by making the exponents of $A_0, B_0, C_0 \ldots .$ the same in the above formula as in the already known dimensional formula.

Sometimes this is not possible. Then we know that there can be no relation whatever between the variables we have chosen and p. If just one single set of exponents $\alpha, \beta, \gamma, \ldots .$ makes the formulae agree, then we conclude that if a relation such as we have assumed does exist, it necessarily has the form we have found. But this by no means implies that the relation does exist. For instance, one might assume that the height H of a human being who had grown up on the moon would depend upon the gravitational acceleration, g, there, and the time he had taken to grow, t. Then

$$H = \text{const. } gt^2,$$

as may easily be found by the usual procedure. But this is undoubtedly a wrong answer. If H is the distance a stone drops, starting from rest, in a time t, however, it is a correct answer physically. Frequently, however, there is more than one set of exponents which will make both dimensional formulae equal. Then an arbitrary function of arguments of zero dimensions finds its way into the answer. An example may, perhaps, be better than a formal demonstration. Let us consider how the distance H travelled by a falling body varies with the time t if the acceleration due to gravity is g and the starting velocity is v. We can, in this case, choose the ordinary mechanical units of mass, length and time for our $a_0, b_0, c_0 \ldots$ units. Then we assume

$$H = g^\alpha \, v^\beta \, t^\gamma \, \phi$$

with the dimensional relation

$$[L] = [LT^{-2}]^\alpha \, . \, [LT^{-1}]^\beta \, [T]^\gamma.$$

Now we can make a dimensionless product out of the variables g, v, t, namely, v/gt. Consequently we know that our formula will finally be

$$H = g^\alpha \, v^\beta \, t^\gamma \, \phi(v/gt).$$

Considering only the dimensional part for the moment, however, we find by equating coefficients that

$$\alpha + \beta = 1,$$
$$2\alpha + \beta - \gamma = 0.$$

Consequently, we have arrived only at a relation between the coefficients and not a unique solution. But the above equations show that

$$\alpha = 1 - \beta,$$
$$\gamma = 2 - \beta,$$

whence

$$[L] = [LT^{-2}] \, . \, [T^2] \, \{[LT^{-1}]/[LT^{-2}] \, . \, [T]\}^\beta,$$

which shows a striking similarity to *

$$H = gt^2 \phi(v/gt).$$

The occurrence of the indeterminable coefficient β is due to the possibility of forming dimensionless products of the variables, and in working out the relations between the indices we infallibly find these dimensionless products. This may readily be seen from the consideration that if a dimensionless product exists, such as

$$[LT^{-1}]/[LT^{-2}] \, . \, [T],$$

the power to which it is raised has no effect whatever on the equality of both sides of the equation—in respect of dimensions—and the exponent is therefore

---

* The normal expression $H = gt^2/2 + vt$ is recognizable as in agreement with this expression if it is written $H = gt^2(1/2 + v/gt)$.

not determinable. But the connection between this indeterminacy and the arbitrary ϕ, which we have already shown to be part of the physical equation, appears to be only incidental. It is probably worth remarking again that the relation obtained by the method of dimensions is not necessarily a sensible one. How sensible the relation is will, in fact, depend upon how sensible we have been in considering and formulating the problem.

§ 4. GENERAL CONCLUSIONS

All that we have required of a physical quantity for it to be expressible by means of the simple formula always adopted in considerations of dimensions is that if two measures of it are equal on one system of units they shall also be equal on any other system. So long as the components which appear in the formula are measured on scales which preserve the same ratio between two measures of a component whatever the units of the scale, any physical quantity can be regarded as having dimensions. There is no restriction whatever to mass, length and time dimensions. In using the " method of dimensions ", for instance, we are at liberty to use any system of units which can be consistently applied to the definition of the quantity we are measuring and the quantities in terms of which we are trying to express it. The intrinsic reality of such quantities or any such esoteric considerations are of no moment.

The only question at issue is : are they necessary in the calculation of the quantity we are interested in? It is completely immaterial, for instance, that the Newtonian constant of gravitation and the inverse-square law are only an approximation; the use of the Newtonian constant in dimensional arguments, about planetary motion for instance, certainly leads to the right answer.

REFERENCES

Bridgman, P. W., 1931. *Dimensional Analysis* (2nd Edn.) (Yale Univ. Press).

Campbell, N. R., 1920. *Physics : The Elements*, Ch. 13 and 14 (Cambridge Univ. Press).

Selwyn, E. W. H., 1935. *Phot. J.* 75, 571-80.

AN ELECTROMETER FOR MEASUREMENT OF VOLTAGE ON SMALL IONIZATION CHAMBERS

By F. T. FARMER, Ph.D.,
Barnato Joel Laboratories, The Middlesex Hospital

Communicated by Prof. S. Russ, 27 *March* 1942

Small condenser ionization chambers, such as the Sievert type, are used extensively in radiological and other work for the measurement of gamma-ray intensities. These have a very small capacity, often as low as 1 $\mu\mu$F. or less, and the measurement of voltage drop due to exposure necessitates the use of a low-capacity electrometer. A fibre instrument such as the Wulf or Lindemann is usually employed for this purpose.

The expense of such an instrument, and the fact that a somewhat elaborate optical system is needed if an external scale reading is to be provided, suggested that some simpler system was worth investigating, particularly for routine measuring purposes. An electrometer valve may be used as the basis of such an arrangement, and it has been found that the device described below, which gives a direct reading on the dial of a meter as a measure of chamber voltage, meets the requirements. The accuracy is as high as that of the electrostatic instruments, and the more compact and robust nature of the apparatus is an advantage also in certain cases, such as where measurements in a hospital radium department are required.

Essentially, the highly insulated grid of the valve takes the place of the electrometer suspension fibre, so that, if the internal electrode of the chamber is connected directly to this while the outer electrode is at a fixed potential, the current through the valve provides a measure of the voltage on the chamber.

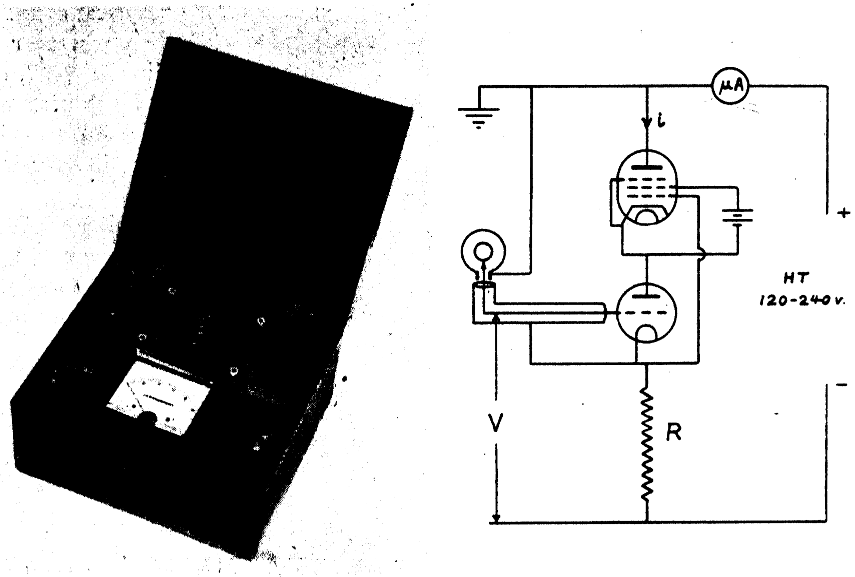


Figure 1.

The voltages which are to be measured are, however, of the order of 100 to 200 volts, so that means must clearly be provided either for reducing the effective voltage on the chamber to that which can be embraced by the grid base of the valve, or for extending by some method the working range of the valve itself. The former could be readily achieved by connecting a fixed capacity in parallel with the chamber, so that the charge of the latter is shared with this, and the potential so reduced to any desired value. The second method, however, has much more important possibilities, since it admits the use of a large degree of negative feed-back, and this can be put to advantage to render the whole measuring system relatively independent of supply voltages as well as of variations in the valve characteristics. The considerable precautions which would be necessary, if the former method were used, to ensure constancy of voltage to anode and filament may thus be done away with, and, as will be seen from the calibration curve (figure 2), even a change from 120 to 240 volts in the high-tension battery does not alter the instrument reading by more than 1 %. Similar considerations apply to the filament supply, and an arrangement which requires the minimum of adjustment for precise measurement is thus rendered practicable.

The electrometer valve must be operated with an anode voltage not exceeding 4 or 5 volts, and, therefore, it cannot be employed simply with a cathode resistance to provide feed-back, on account of the large voltages which would appear at its anode. To meet this difficulty, a second valve of high impedance is placed in series with it, as shown in figure 1. The bias potential for the grid of this valve is obtained directly from the voltage drop in the electrometer valve itself, and in consequence the potential difference between anode and filament of the latter is kept to its required low value. The encumbrance of the second valve is not great, since for the small current which it has to pass, only a few volts are required to supply its screen, and its filament may be operated from the same source as the electrometer.*

The two valves together function in the manner of a " cathode follower ", and it is clear that the relationship between grid potential, V, relative to the negative terminal of the battery, and the reading, i, of the microammeter, is given very nearly by

$$i = V/R,$$

where R is the cathode resistance. The characteristics of the valves and batteries only enter to a second order, and approximate linearity of the overall calibration comes about also through these considerations. A cathode resistance of the order of 1 megohm, allowing a maximum current of about 100 micro-amperes to flow through the instrument, is employed.

The inherent capacity of the grid and its connecting lead to the chamber needs to be taken into account, since the charge to be measured is shared with this. With the valve used, a Marconi E.T.1, it amounted to 3 $\mu\mu$F., which is about three times the chamber capacity, and if no steps were taken to counteract it, it would mean that only a quarter of the voltage on the chamber would appear at the grid, and the negative feed-back could not be made as effective as is desired. To overcome this difficulty, the valve and grid lead were surrounded by a screen connected to the electrometer filament, the potential of which follows very closely that of the grid itself; the capacity to earth, therefore, of the grid and lead is reduced to a very small fraction of its normal value, and almost the whole voltage on the chamber appears at the grid as is desired.

In using the instrument, the chamber must be charged negatively relatively to its case, so that when it is applied to the electrometer, the sudden change in the grid potential is in a negative, not a positive, direction. This is necessary, since, in the transient condition, the cathode potential of the valve lags very slightly behind that of the grid, owing to the appreciable capacity to earth of the filament battery, and if the grid is allowed to assume a more positive potential than its equilibrium value, grid current flows, and the charge is incorrectly registered. A momentary change to a more negative potential, however, causes no flow of current, and, therefore, no disturbance to the working.

---

* A 2-volt valve (type 15) with indirectly heated cathode allows this simplification with battery-operated filaments.

In order that the reading of the meter should remain steady after contact with the chamber is made, it is essential that the grid current in the static working condition should be extremely small. The valve mentioned above has been found very satisfactory from this point of view, no perceptible drift of the meter occurring in the time required for taking readings.

Changes in the supply voltages to the instrument are, as pointed out, immaterial unless they are severe, so that no provision is necessary for adjusting high-tension or low-tension voltages. The small current taken from the high-tension battery means, moreover, that there is little tendency for it to run down except at long intervals. A magnetically operated switch is incorporated for "earthing" the grid of the electrometer valve before it is brought in contact with the chamber, so that it should start from the same potential in each measurement.

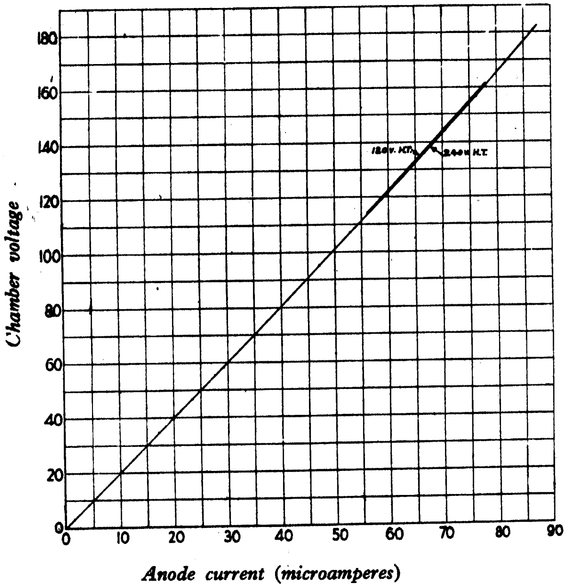


Figure 2. Voltage calibration for Sievert chamber, type (1–1).

A feature of this type of apparatus worth mentioning is that it is not harmed by applying greatly excessive charges to the electrometer valve grid, as may occur when the chambers are used in experimental work. The possibility of damage to a sensitive fibre electrometer through such causes is a factor which has constantly to be borne in mind when such instruments are used.

I am very greatly indebted to Professor S. Russ for providing all facilities necessary for this work; and also to Miss D. Clephan for the photograph reproduced herewith.

A HIGH-VOLTAGE APPARATUS FOR ATOMIC DISINTEGRATION EXPERIMENTS

By J. D. CRAGGS, Ph.D., M.Sc., A. Inst.P.,
Metropolitan-Vickers Electrical Co., Ltd., Manchester

MS. received 2 *February* 1942 ; *in revised form* 13 *April* 1942

ABSTRACT. A simple apparatus for performing experiments in nuclear physics has been developed to the stage where beam currents of 100 to 200 μa. at 600 kv. have been obtained. Experiments with the acceleration tube have led to some interesting results ; the presence of large currents of secondary electrons has been confirmed and their effect on the running of the apparatus is discussed in some detail. It was hoped to carry out experiments on the installation of different forms of limiting diaphragm in the tube, but lack of time and the wish to obtain some results with proton bombardment prevented our doing so.

Some experiments with nuclear γ-rays are briefly described. The results are not accurate, but they have indicated the practicability of continuing such work.

The performance of the tube compares very favourably with other installations using electrostatic generators for the supply of high-tension current. The output from the ion source is capable of being raised almost by a factor of 10, so that if a more powerful generator had been available, target currents of the order 500 μa. should have been easily obtained with a heat dissipation higher probably by a factor 4 than that encountered in the above work. This presents only a minor technical problem.

§ 1. INTRODUCTION

Since the pioneer work of Cockcroft and collaborators, Tuve, and others, there have been many investigations in the field of artificial atomic disintegrations, and nuclear structure has now become a specialized branch of physics, with its own established techniques. One of the principal requirements in the production of artificial nuclear reactions is an apparatus for accelerating charged particles (usually protons or deuterons) to high speeds, and in large numbers. Disintegrations can, of course, be produced with fairly slow ions ($\sim$50 kv.), but the reaction efficiencies are low, as penetration of the nuclear barriers is a probability effect which increases rapidly with increasing speed of the bombarding particles. Hence, until voltages comparable with the barrier heights are attained, the apparatus should work at as high a voltage as possible. Further, since the disintegration output will vary linearly with beam current, the effect of a low current may be offset by a smaller relative increase in voltage which, in general, is more easily obtained. The limit, of course, is set by the size of building and the increasing cost of the apparatus.

In designing the present apparatus in outline, emphasis was also laid on the need for flexibility—for instance, the use of a Wilson chamber at the target end of the tube instead of Geiger counter or ionization chamber detectors. It is also important to facilitate the irradiation of samples with the greatest possible efficiency, i.e., to use a large target solid-angle, as many induced reactions are of low efficiency—for example, production of photo-nuclear effects by γ-ray bombardment.

§ 2. REQUIREMENTS FOR THE PRESENT APPARATUS

It was decided to build a neutron generator on modest lines with an output voltage of 700 kv. and a beam current of 200 to 500 μa., as this seemed an economical proposition. Available space was limited by the use of a room approximately 16 ft. wide × 23 ft. high in King's College, London. The sideways clearance was the controlling factor. During the development period in Manchester, therefore, these figures were carefully considered. Owing to

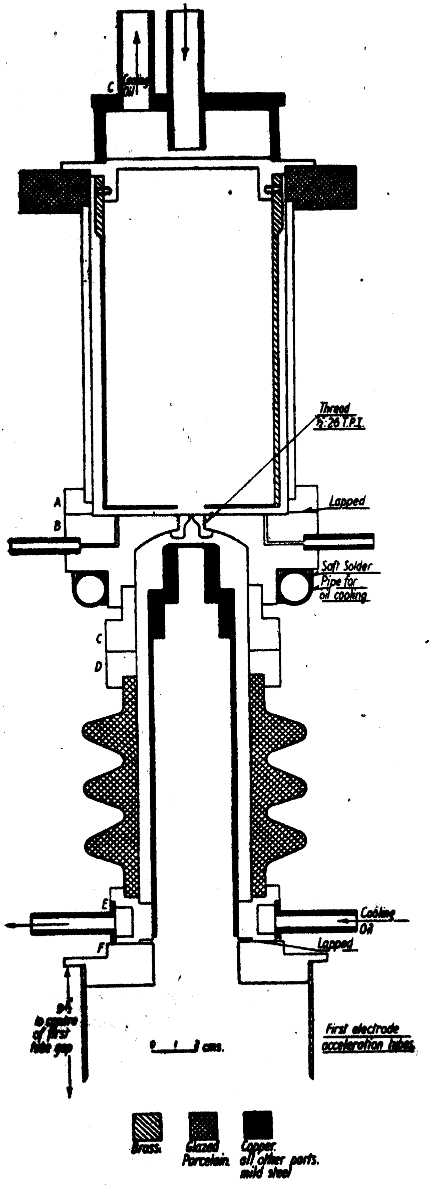


Figure 1. The ion source and focusing electrode assembly. To avoid complicating the diagram the following details are omitted: (1) Bakelite clamping bolts and nuts for ion source with brackets on B and C; (2) centring screws bearing on A with brackets on B; (3) centring screws bearing on C with brackets on D; (4) bakelite clamping bolts with brackets on B and F; (5) centring screws bearing on E with brackets on F; (6) steel clamping bolts with brackets on F and first plate of acceleration tube. Interior surfaces of ion source are polished.

the present circumstances, the apparatus has not been installed in London, and the preliminary programme of work has been completed in the Metropolitan-Vickers Works at Trafford Park, Manchester.

The apparatus could have been designed with the target at a high, not earth, potential. Before discussing the reasons for adopting the latter arrangement, a brief description of the essential components of the equipment will be given. It is necessary, first, to produce an intense beam of positive ions—a cold-cathode discharge tube was used in our case (Oliphant and Rutherford, 1933)—and then to focus the beam into the main accelerating tube, after which the diffusion and scattering losses should be small, so that the ions proceed to the target. It is possible to use the first one or two stages of the cascade tube for focusing, but the interposition of a separate focusing stage is desirable. The electrode assembly at the inlet end of the acceleration tube is shown in figure 1. The source and focusing will, it is hoped, be described in separate publications. If the source were arranged to be at or near earth potential, manipulation of the controls would be simplified. Further, the pumping equipment, which is necessarily at earth potential, would be utilized to the best advantage, because the gas flowing from the source through the canal into the acceleration tube would be removed near the canal. However, the target would be at a potential of 700 kv. with respect to earth, and many measurements, e.g. counting experiments or the use of a Wilson chamber, would be complicated. On the other hand, if the target were earthed and the source at high potential, as in our arrangement, the pumps would be somewhat choked by the acceleration tube, and the power supplies for source and focusing would have to be remotely controlled. The latter scheme, on account of the large high-tension structures, is more expensive but, on the whole, more convenient, because the target end of the tube is of more importance in running than the source, which, once adjusted, rarely needs attention. The earthed-source arrangement is seldom encountered.

§ 3. DETAILS OF CONSTRUCTION

The following is a brief description of the method of construction of the high-tension electrodes. Wooden frames (2 in. × 2 in. members) were bolted to channel-iron base structures (3 in. × 2 in. × $\frac{3}{8}$ in.), which in turn were secured to the bakelite cylinders. Finally, the structures were covered with $^{1}/_{32}$-in. sheet steel, the end caps being mounted first, with steel supporting bars, 1 in. × $\frac{1}{8}$ in., attached to the wooden framework. The final appearance of the apparatus is seen in figure 2. The van de Graaff generator is 112 in. high, its terminal being 57 in. in diameter. The height of the electrodes is about 4 ft. 6 in. The bakelite cylinders are 75 in. high, 7$\frac{3}{4}$ in. internal diameter and $\frac{3}{4}$ in. thick. In figure 2 metre scales are shown, fixed to the right-hand front support of the large electrode and to the front ground girder.

The design of the frames was carefully considered. The total weight carried by the two sets of bakelite supports was about 1$\frac{1}{2}$ tons. The bakelite cylinders were adequately strong: the framework of the larger electrode was the most highly stressed part. Even so, it was estimated that the safety factor was about 8, which was considered adequate. Vibrations are only appreciable when the driving motor, on slowing down after a run, forces the electrode into resonance, but the effect is not serious.

It was necessary to determine satisfactory values for the electrode radii in order to minimize corona losses, as distinct from increasing the spark-over voltages. The former were important because of the limited current output of the van de Graaff generator. From Peek's (1929) data for cylinder/cylinder gaps it is possible to deduce the maximum gradients obtaining in cylinder/plane gaps: the latter are relevant to a study of the practical case of our high-voltage electrodes. Taking a total voltage of 850 kv. for a gap of 7 ft., and cylinders of 6 in. radius, the maximum gradient is approximately 29 kv./cm., which is near the breakdown value for normal air, viz. 30 kv./cm. As we did not expect to exceed 700 kv. in our earlier work, 8-in. radii were considered sufficient.

The minimum clearance to the laboratory walls (see figure 2) was about 6 ft. to a projecting girder: although this spacing is approximately the same as that for the electrode/ground, the latter is of less importance, because the floor provides a large conducting surface which acts as a screen. The driving motor and belt seem to exert no appreciable influence. The electrode coverings (figure 2) were secured to a hoop system of steel strips with screws, the heads of which projected outwards. These have never brushed appreciably, because of the screening effect of the large metal surfaces.

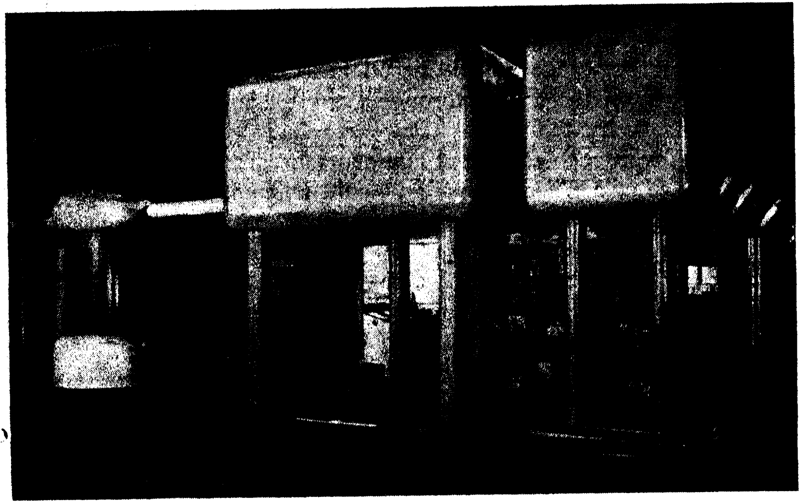


Figure 2.

Figure 3 shows corona losses for the complete apparatus. The beneficial effect of thin pressboard insulating screens interposed between the wall and the high-voltage electrode is to be noted. This technique has been used in many laboratories (e.g. by Bouwers at Eindhoven and at Cambridge). The effects of introducing such barriers into asymmetric air and oil gaps were described many years ago (Goodlet, Edwards and Perry, 1931). The breakdown voltages of gaps which do not form corona before spark-over are in general decreased by the interposition of a barrier of higher S.I.C. than the initial medium. However, in the presence of even small corona currents, the reverse holds, i.e. spark-over voltages are increased, due, presumably, to the barrier becoming charged to a potential near, and of the same sign as, the high-potential electrode, and so acting as a screen. The tentative theory is supported by the fact that the efficacy of the barrier is somewhat increased as it is made to approach the high-potential electrode. There seems to be an optimum screen position for most gaps (Goodlet, Edwards and Perry, *loc. cit.*). The total available output from our van de Graaff generator was about 1·5 to 1·8 ma. positive excitation, at about 700 kv.

The arrangement of equipment in the electrodes is shown diagrammatically in figure 4. The lead to the van de Graaff generator terminal is on the extreme right-hand side. (See also figure 2.) In the larger electrode there are two D.C. sets (+50 kv. and −75 kv. maximum) each " earthed " to the frame, which is at the potential of the ion source (discharge-tube) cathode. The focusing electrode (see figure 1) and the first cylinder in the cascade acceleration tube are connected to the small electrode. The arrangement whereby the cathode is at the potential of the common terminal of the D.C. sets was adopted for technical reasons, e.g. the avoidance of D.C. sets giving voltages >75 kv., which were considered high enough for focusing the ion beam, as the discharge tube was intended to operate at ~30 kv. with a possible 50 kv. maximum. The D.C. sets are remotely controlled by bakelite rods, chain-and-sprocket driven at the control end, and secured to the auto-transformers controlling the high-tension transformer primaries in the electrode. On-off contactors (fitted with overload relays, fuses, etc.) are operated by strings. The ion-source anode is cooled by special oil of low viscosity flowing in insulating tubes through the upper busbar (figure 4) and impelled by a small centrifugal pump driven off the 9·5 kVA. alternator in the electrode, which is driven by a 12·5-H.P. motor at earth potential (shown in figure 2): a cooling system, fan and radiator, for the oil is similarly driven. The source and focusing voltages are read from milliammeters, measuring the currents flowing through special high resistances in parallel with the electrodes (see § 6), mounted on a central panel with signal lights, and visible through an aperture cut in the electrode. The small electrode

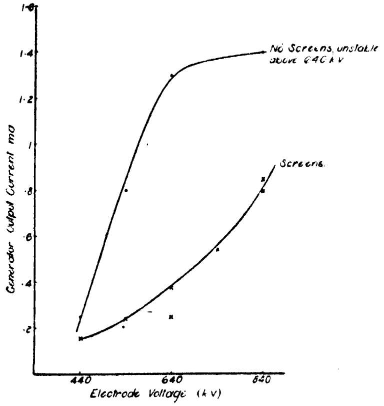


Figure 3. The effect of insulating screens on corona lines from high-voltage apparatus.

contains, when required, a magnet for separation of unwanted ions from the beam. Power is supplied by a D.C. generator driven by a 2·5-H.P. motor and controlled rheostatically. Usually the magnet was not required, and was then removed to simplify control. Alternatively, the magnet may be mounted at the earthed end of the tube.

The ripple on the focusing and source voltages was $\not>$ 5 %. The only point of interest in the D.C. sets was the horizontal rectifier valve mounting, adopted to reduce the electrode height to a minimum, and incorporating rubber pads to damp out vibrations.

The electrode system and mounting of the cascade tube are normal (see § 7). However, in order to economize in head-room, the tube was inclined to the vertical, a procedure which increases the difficulty of supporting the tube, adjusting it, and making the joints. After traversing the acceleration tube, the ion beam passes through the tee-piece connected to the diffusion pumps, a 1¼-in. bore conical tap, and finally strikes the target, attached to a copper holder. The latter is of conventional design, cut away at 45° to the axis and water-cooled. The techniques adopted for using different targets are described below. Figure 2 shows the corona shields on the tube, the diameters of which were decided upon after calculations similar to those described in some detail for the large electrode.

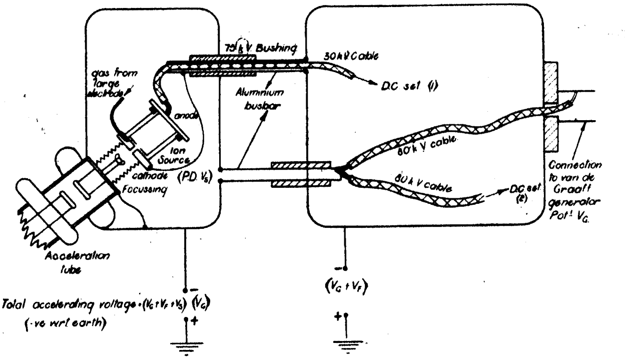


Figure 4. Potential distribution in neutron-generator electrode.

V_F=focusing P.D. ∼ 10/70 kv.
V_S=source P.D. ∼ 10/20 kv.
V_G=van de Graaff potential (∼ 600 kv.).

The van de Graaff generator is controlled from a separate cubicle, and the voltage applied to the large electrode is measured with a generating voltmeter of simplified design. The latter is calibrated with a sphere gap checked in the usual way by observations on gamma-ray resonances, which were supplemented by tests carried out with a set of wire-wound high resistances: we estimate that the error of voltage measurement is about ± 2 %. This was considered satisfactory for preliminary work, and the need for extra stabilization has never arisen.

§4. PUMPING SYSTEM

Because of the appreciable quantity of gas, about 300 c.c. of hydrogen per second at 0·1 to 0·2 mm. Hg pressure, flowing from the ion source, high-speed pumps must be used to give low pressures in the acceleration tube. The mean free path for any likely collision processes should be at least equal to the tube length. Rough extrapolations for ionization and *Umladung* losses in hydrogen (Gerthsen, 1930; Meyer, 1937) indicate mean free paths, at 10^{-4} mm. Hg pressure, for 200-kv. protons to be $\sim$1 m. and $\sim$30 m. respectively. The tube length was about 80 in. (2 m.), to give suitable external gap lengths for 700 kv. operating voltage. Hence the tube pressure should be $\sim 10^{-4}$ mm. Hg, and preferably 10^{-5} mm. Hg. This value agrees reasonably with published data (Bouwers *et al.*, 1937; Lauritsen *et al.*, 1941), and from the data given above necessitates a pumping speed of 200 to 1000 l./sec.

The pumping set, as finally developed, consisted of a 6-in. diameter brass pump, based on a design by Crane, in series with a 3-in. pump, a Metropolitan-Vickers type 02 and a Cenco-Megavac rotary pump. The diffusion pumps were mounted in an angle-iron frame adjustable for height with corner bolts. The main intake, connected to the acceleration tube, is 6 in. in diameter and is bolted down to the 6-in. pump, using a lead or rubber gasket. The pump speed at the entrance to the acceleration tube was about 300 l./sec. for hydrogen. Subsequent experience has shown this performance to be adequate.

§5. DETAILS OF MISCELLANEOUS ACCESSORIES

Before describing the experiments carried out with the complete apparatus, it was decided to include a section giving details of some of the component parts of the apparatus, where they might be considered useful to other experimenters.

The resistance voltmeters used for measurement of source and focusing voltages are of interest (see § 4 for mention of their installation, etc.). An accuracy of ± 5% was considered sufficient. Further, it was required to use robust instruments which could be easily read at distances of about 10 ft.

Specially constructed wire-wound resistances of considerable accuracy, up to about 150 megohms in value, have been used regularly in this laboratory for many years, but were considered unnecessarily exact for this work. We used chains of Morganite 1 or 2 watt resistances arranged in a helix round bakelite supports, and immersed in transformer oil to give adequate cooling and freedom from corona losses.

Galvanometers or electrostatic voltmeters were rejected as indicators in favour of 1-ma. meters. The usual technique adopted in these laboratories is to run such resistances at much less than their rated dissipation because of their high negative temperature coefficients, but we decided, size being a very important consideration, to use currents of the same order as the nominal ratings. For the 50-kv. source voltmeter, 50 1-megohm, 1-watt (1-ma.) resistances were used, as it was realized that 30 kv. would be rarely exceeded: for the 75-kv. focus stage voltmeter, the resistance chain consisted of 15 1-megohm, 2-watt (1·4-ma.) resistances. The spiral assemblies, with brass end plates, fitted closely inside 8-in. diameter bakelite containers and were 9½ in. and 15 in. long respectively.

The nominal 50-megohm unit (50·5 megohms at 20° c.) changed by <3 % at 0·5 ma. in the steady condition. At 0·75 ma., when stability was reached in about 5 hours, the resistance change was about -6 % for a temperature rise of 19 to 20° c. The resistance was linear and free from appreciable hysteresis effects. It was shown by special experiments that the resistance changes were due to the Joule heating and not to any voltage effect. The effect of oil immersion on resistance was shown to be negligible. Further experiments performed after a lapse of two months showed the resistance change to be <1 %.

A test with the nominal 75-megohm resistance (actually 78·6 megohms at 20° c.) gave -10 % change for 1 ma. current, which was considered excessive. There are several standard methods for reducing such effects, e.g. thermostatic control or the use of a potentiometer arrangement whereby the voltage measured is a fraction (independent of temperature changes if the heating of all the resistance units is equal) of the total. A method was devised which seems to be simpler than those just described: it consists of immersing a metallic resistance R_2 in the oil container, so that R_2 reached approximately the same temperature as the main resistance R_1, and applying a shunt R_3 of negligible temperature coefficient across the meter and R_2 in series. R_2 increases as R_1 decreases, so giving a compensation effect. The meter resistance should be much less than R_2 to avoid swamping the compensation effect, and R_2 should equal R_3 for greatest sensitivity. Typical values were:—meter resistance $=40$ ohms, $R_3 \simeq R_2 \simeq 1000$ ohms. Without attempting to obtain the best conditions, i.e. equality of positive and negative temperature coefficients, the effective change in R_1 was reduced to the acceptable value of -4 % for 30° c. rise in temperature, i.e. at 1 ma. current, using a copper coil for R_2.

The measurement of the van de Graaff terminal voltage presents some points of interest. The method used should involve as small a current consumption as possible, and should give continuous readings with stability and ~ 2 % accuracy. The last requirements eliminate sphere gaps, and a generating voltmeter (Kirkpatrick, 1932) becomes the obvious choice. The operating principle is, briefly, as follows. An insulated plate is alternately exposed to and shielded from the high-voltage electrode whose potential we require to measure when shielded, e.g. by a rotating vane system, the pick-up plate loses its charge through a resistance. If the latter be connected to a suitable galvanometer or valve amplifier, a direct reading of voltage is obtained. The instrument is not absolute, although it becomes so in carefully defined geometrical conditions, and is usually calibrated with a spark gap and with γ-ray resonances where the voltmeter is used with atomic disintegration apparatus. A simple single-stage valve amplifier was used (10 to 20 megohm grid leak, depending on required sensitivity) giving 0 to 500 kv. linearly above a small initial exponential part of the curve, with a 0 to 1 ma. meter. A $\frac{1}{8}$-H.P. synchronous motor (1500 r.p.m.) was used to drive the rotor.

It is not possible for the induced changes on the pick-up plate to be measured directly with a D.C. instrument, because the direction of the current changes when the pick-up plate is covered, i.e. twice every revolution. If C is the capacity of the H.T. electrode/pick-up plate assembly with voltage V applied to the electrode, then if Q be the charge at any instant,

$$\frac{dQ}{dt}=i=V\frac{dC}{dt}.$$

dC/dt changes sign when the pick-up plate is completely covered, so that the mean value of dC/dt is zero.

The R.M.S. voltage (E) across the load resistance (R) is shown (van Atta *et al.*, 1936) to be

$$E=3\cdot93\times10^{-10}\ F.A.S.R.\ \text{millivolts},$$

where S = rotor speed (r.p.s.), A = effective rotor area, and F = field strength at the rotor in volts/cm. In one case $S=25$, $A\simeq160$, $F\simeq4\times10^3$, $R=10^7$, in the appropriate units. Hence $E\simeq60$ volts. This is higher than the value obtained in practice by a factor ~10, probably because F is much lower than the value taken here, which equals the average gradient from the generating voltmeter to the high-potential electrode. It is clear that the method is capable of great sensitivity. Over an operation period of about six months the voltmeter has shown great stability.

As an even simpler alternative to the above, we attempted to make high liquid resistances ($\sim10^{10}\,\omega$) using special liquid mixtures in Pyrex containers. The best resistances showed steady changes of about 1 %/day, due probably to differential evaporation of the rather complex fluid. Such resistances would have been satisfactory if re-filled and calibrated once a week, but cannot be recommended for precision work. In view of the interest attached to this problem a separate publication has been prepared and is now in the press.

§6. EXPERIMENTS ON THE ACCELERATION TUBE

The tube construction may be seen in figure 2. The cylinders ($^1/_{16}$ in. thick brass) are $9\frac{3}{8}$ in. long, 4 in. internal diameter, and are provided with steps to prevent sliding through the steel end plates which project beyond the glass (6 in. internal diameter, $^3/_{16}$ in. thickness, and 10 in. length) cylinders. The brass electrodes are maintained in alignment because of their driving fit in the end plates, the bearing surface being $\frac{3}{8}$ in. long. The stress distributors mounted on the ends of the electrodes, inside the accelerator tube, were formed from copper tube of $\frac{3}{4}$ in. outside diameter, arranged to give an inter-electrode spacing of $\frac{1}{2}$ in. This last setting is not critical except, possibly, in the first stage of the tube, as the focusing is obtained largely by the use of the primary focusing electrode (figure 1). The glass cylinders are clamped between the end plates by the use of $\frac{1}{2}$-in. diameter Tufnol rods and nuts. All vacuum joints on the tube were made of Apiezon Q compound. The external stress distributors (figure 2) were graded from 7 in. to $1\frac{1}{8}$ in. diameter, the latter at the earthed (target) end of the tube. External spark-over for a single gap with $1\frac{1}{8}$-in. diameter shields was 140 to 150 kv. The alignment of the tube was excellent, as the glass cylinders had been carefully ground to close limits. Lapped flanges with centring screws are provided at the inlet and exit ends of the tube. The ion beam passed through a 12·2 mm.-diameter collimating aperture before passing through the tap to the target.

The first experiments showed the existence of large currents of secondary electrons produced by the relatively small number of stray positive ions in the tube. This phenomenon had been expected, on the basis of other work with high-speed positive ions (e.g. Hill *et al.*, 1939). For instance, at 240 kv. (negligible corona) some experiments showed a total tube current of 0·6 ma. when the target current was 20 μa. and the entrant positive-ion current ~200 μa.

Clearly this problem is of great importance where the largest possible target current is required with a limited available output current from the van de Graaff generator. Some time was spent in investigating the tube behaviour, as the published data seemed negligible despite the inevitable appearance of the effect in most of the tubes described (e.g. Gray *et al.*, 1940). The most obvious remedy is to prevent the secondary electrons from leaving the regions in which they are produced. In a well focused tube, most of the secondaries are produced at the earthed end of the tube, where beam spreading may become appreciable. An interesting consequence of the electrons' travel up the tube is the production of more secondaries by multiplication at the various electrodes, which would modify the potential distribution. It is desirable that the latter should be determined only by the external corona gaps fitted to the electrodes, and not by the internal currents.

To measure the total x-ray output, a calibrated gold-leaf electroscope was used (5·6 divs./min. ≡ 10^{-5} r./sec.) of dimensions 3·3 × 3·3 × 5 cm. It was shown first that all the increased leakage was due to radiation formed in the presence of a beam: switching off the latter with maintained voltage always reduced the discharge to a very small value (natural leakage rate). The electroscope was set up 12 ft. from the tube opposite the high potential end, and 4 ft. below the level of the first tube gap. The results are shown summarized in table 1. The brass electroscope case was 1 mm. thick.

Table 1

Experimental condition	Rate of electroscope discharge (divs./min.)
Natural leakage, etc.	0·08
γ rays from 0·5 mgm. Ra 3″ from centre of electroscope	1·3
16 μa. beam ; 240 kv. ; tube current 0·6 ma.	1·2
20 ,, ; 240 kv. ; ,, ,, 0·6 ,,	2·0
4·5 ,, ; 420 kv. ; ,, ,, 0·3 ,,	5·2
3·7 ,, ; 300 kv. ; ,, ,, 0·3 ,,	1·7
Electroscope mounted 3″ from target	
3·7 μa. beam ; 300 kv. ; tube current 0·3 ma.	1·7

The safe dose is taken as 10^{-5} r./sec. (Kaye *et al.*, 1939) for 200 working hours a month. This corresponds to an electroscope discharge rate of 5·6 div./min. With the tube running at its maximum dissipation, say 160 μa. beam at 600 kv. with a total tube current of 1·2 ma., the safe discharge would be considerably exceeded unless the observers were shielded. No measurements were taken in these conditions.

Some performance figures for the tube without the separate focusing stage (see figure 1) are:—34 μa. target current for total tube current of 0·45 ma.

(positive and negative current) at 310 kv. with the ion source running at 9 kv., 0·5 ma., i.e. total canal-ray beam of $\sim$ 100 μa. The tube efficiency, relative to the collection of positive ions only, is good (probably $\sim$50 % or more). The inclusion of a collimator at the entrance of the tube gave no appreciable improvement.

The establishment of a weak magnetic field ($<$50 gauss) near the pump connection at the base of the tube gave a rise in accelerating voltage, for a constant van de Graaff generator output of 0·4 ma., from 390 to 450 kv., with a simultaneous rise in target current from 20 to 25 μa. Clearly the return current of secondary electrons had been much reduced. Target currents of 50 μa. were easily obtained. The tube voltage was reduced to 100 kv. to reduce the accelerating field for the secondaries, and we then obtained target currents of 8 μa. with no field and 3 μa. with the lowest tube gap shorted to reduce the field still further. In the latter condition 20 μa. were obtained with the field established near the pump connection. The secondary electrons were expected to arise mainly in the pump connection where the beam emerges from the last gap and impinges on a molybdenum diaphragm before entering the target tube.

It must be emphasized that the increase in target currents, as distinct from the rise in tube voltage for given input current when the secondary electron-deflector field is switched on, is largely due to a change in the tube-potential distribution. The voltage rise is due mainly to the decreased drain on the van de Graaff generator, which runs as a constant-current device.

Some results on potential distribution were obtained. A sphere gap comprising copper spheres of 6·25 cm. diameter on threaded shanks was placed across each tube gap in turn and adjusted with a long insulated rod to spark-over occasionally, but regularly, with the tube in operation. It was assumed that the presence of such a device would not affect the potential distribution: in view of the absence of corona discharge from the spheres this is probably justified. The sphere gap was mounted on the opposite side of the tube from the voltage-control corona gaps, which were not altered during a run. It was expected that tube surges would vitiate the results, but as they occurred at random they could usually be detected. The results are shown in figure 5. The total tube potential for (A) was 440 to 445 kv., and the sum of the separately measured gap voltages 448 kv. For (B) the respective figures were 445 and 569 kv. (surges might be expected to be serious here because of the appreciable beam current). For (C) the readings were 445 and 437 kv. The presence of a beam clearly altered the potential distribution, which was nearly linear in the absence of a beam. Figure 6 shows the voltages for each separate gap. The increased gradient at the low-potential end of the tube is the most prominent feature of the results. For all these tests, 1 in. long copper wires were attached to each electrode, pointing towards the positive electrode of each pair. This arrangement gives the most stable corona condition (confirmed by van Atta and Nowak, 1941). The corona wire on the earthed electrode brushed violently when large beam-currents were used, indicating (see figures 5 and 6) a large gap voltage. The total tube-current, ignoring small differences in corona current to earth from the various electrodes, will be constant at all points, i.e. a low internal electrode-to-electrode current necessitates a high external corona current for balance.

Several explanations of the effect of the beam on voltage distribution are possible. For instance, the cascading of secondary emission, i.e. successive multiplication at the different electrodes, would tend to decrease the gradient at the high-potential end of the tube. Again, the incoming ion-beam may lose its slow components to the upper electrodes: this suggestion is rendered unlikely by the known large effect of small deflecting fields at the pump connection, and the fact that inclusion of a collimator below the focusing gap had little effect.

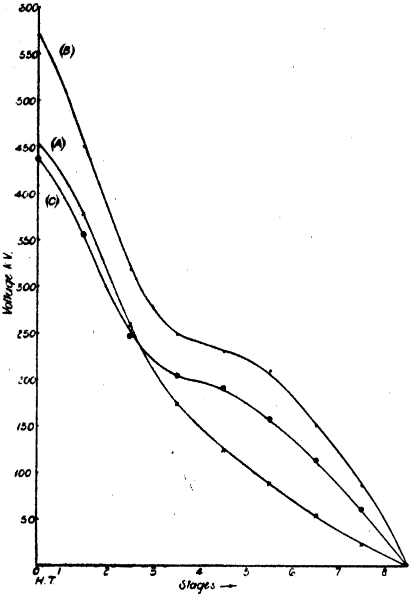


Figure 5. Potential distribution in cascade acceleration tube.

(A) No beam.
(B) 17·5 μa. beam.
(C) 5 μa. beam.

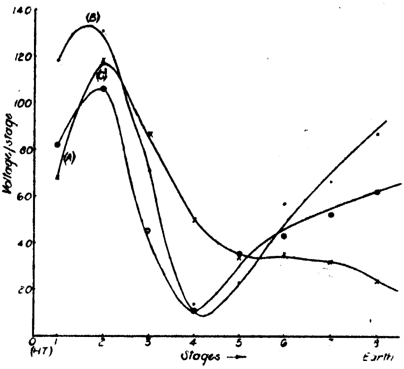


Figure 6. Variation of gap voltage along cascade acceleration tube.

(A) No beam.
(B) 17·5 μa. beam.
(C) 5 μa. beam.

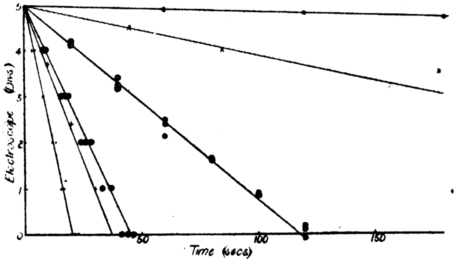


Figure 7. X-radiation from cascade tube.

* Natural leak.
· No absorber.
⊙ 0″·032 Pb.
× 0″·070 Pb.
⊕ 0″·0625 brass + 0″·25 glass.
+ 0·2 mgm. Ra. 1″ from centre of electroscope.
0·98 ma. van de Graaff output.
520 kv. 2·5 μa.
Ion source :—5 kv., 1ma.

The first tube electrode (highest potential) always became hot during operation, so it was clear that the secondaries had attained considerable energy. Probably many of the secondaries ($\sim$60 %) arose at the earthed end of the tube and travelled through several gaps before striking any of the electrodes. In view of the interest and importance attached to the problem, we decided to measure the hardness of the γ radiation from the tube, which should give information on the speed of the secondary electrons and thus on their region of formation. It was not expected that clearly defined results would be obtained, because of the continuous nature of the secondary-electron effects in the tube.

The electroscope previously mentioned was used for these tests. Some of the results are shown in figure 7. We attempted to correct for the absorption in the electrodes and glass tube walls by finding the effect of enclosing the electroscope in a similar electrode and glass tube, the equivalent lead thickness of which was taken as $0{\cdot}3 \pm 0{\cdot}1$ mm. ($\simeq 2{\cdot}75$ mm. copper).

In order to determine the hardness of the rays from the absorption data we have

$$\frac{I_1}{I_2} = \frac{e^{-\mu x}}{e^{-\mu y}} \quad \text{or} \quad \log_e \frac{I_1}{I_2} = \mu(y-x),$$

where I_1 and I_2 are the emergent intensities of radiation for equal incident intensity and for absorber thicknesses x and y respectively. It is assumed for convenience that μ, the linear absorption coefficient, is independent of x or y, i.e. that the radiation is monochromatic. This assumption is false in our case, but the results shown in table 2 indicate that the errors so introduced are not serious.

Table 2

Thickness of Pb absorber (mm.) excluding zero correction	μ	Equivalent speed of electrons (kv.)
3·43 to 1·78	10·4	250
1·78 to 0·81	7·8	280
1·78 to 0·81	13·5	230
1·78 to 0·3	16·5	210
1·78 to 0·3	18·5	190

The voltages corresponding to different values of μ are taken from Wintz and Rump (1931). The first reading is probably the most, and the last the least, accurate because of uncertainties attached to the zero correction. The deflection magnet mentioned above was not used in these experiments. If the radiation is homogeneous, i.e. if the differences in column 3 of table 2 are due entirely to experimental errors, the bulk of the secondaries must arise at the middle of the tube.

As mentioned above, it is remarkable that so little attention has been directed to the problem of secondary currents in acceleration tubes. For D.C. generators using thermionic valves, the increased tube current is usually not a disadvantage except for the heating and the x radiation emitted, but with electrostatic generators the problem is of prime importance (see § 2). Limiting diaphragms in the electrodes have been used since the early work of Cockroft, but it is doubtful whether they were of great value. We made the necessary modified electrodes but were not able to test them.

Some further tests on focusing were carried out, to reduce the beam spread, which in turn would decrease the amount of secondary emission. Decreasing the corona gap on the lowest stages of the tube from 5 in. (point/plane) to $2\frac{1}{2}$ in. (point/point) decreased the target current in a particular experiment from 70 to 22 μa. A remote-controlled corona gap was fitted to the highest potential stage but had only a small effect on focusing. On the second stage the effect was not observable. The separate focusing stage (figure 1) has a large effect on the tube performance, and by the use of it the target current may be almost doubled.

Finally, the use of a magnetic field of several hundred gauss in the tube near the pump connection resulted in a fall in the requisite input current to give 60 μa. target current at 450 kv. from 1.0 ma. with no field to 0·6 ma. with field. The effect of the secondary emission on the target current was negligible, i.e. about 10 %, shown by a separate experiment. The target was effectively shielded from the electrostatic field of the acceleration tube by the interposition of the tap and target tube, equivalent to a field-free region about 12 in. to 18 in. long and 2 in. diameter. A magnet was specially built for this work, and was installed below the earthed end of the tube. The best tube performance thus obtained was 160 μa. target current for 1·8 ma. van de Graaff current at 500 to 600 kv. The focusing voltage was 10 kv. for a source voltage of 6 kv. at 2 ma. current. The total positive-ion current entering the accelerating tube was about 200 μa. The tube-focusing efficiency is high and the secondary currents, amounting to about 1·5 ma., allowing for corona losses, are caused by $\sim$50 μa. of lost ions, together with some neutral particles. Hence the magnification due to secondary emission at the electrode, or ionization processes in the gas of the tube, is about 30.

The beam current in the tube, with hydrogen, consists of about 50 % protons and 50 % molecular ions. A separate communication (Craggs, 1942) deals with this matter more fully, but the importance of the data in relation to the formation of secondary currents is easily seen. Measurements of secondary-emission ratios for H_1^+ and H_2^+ (78 to 426 kv.) show (Hill *et al.*, 1939) that values of about 6 and 3 for H_2^+ and H_1^+ ions respectively may be obtained for non-degassed copper surfaces. The molecular ions may be considered for bombardment purposes as two loosely bound atomic particles of half the equivalent H_2^+ voltage. Appreciable currents (Craggs, 1942) of fast neutral particles exist in the tube, so that a magnification of 30 or more may easily be obtained, owing to the almost certain existence of repeated secondary emission in the tube. In general, secondary-emission ratios tend to be approximately 2 for degassed copper surfaces (Allen, 1939), but the argument remains substantially unaffected.

§ 7. EXPERIMENTS IN PROTON BOMBARDMENT OF LIGHT ELEMENTS

It was decided that the first nuclear experiments should be on the γ-ray resonances of light nuclei. These have been investigated many times, but the absolute yields had not been accurately measured when our work was carried out. Since that time some data have appeared (van Allen and Smith, 1941: Streib *et al.*, 1941). Another reason for the value of such experiments is that they provide a ready and accurate means of calibrating the apparatus and checking its general performance. The thin target resonances are often very sharp, e.g. the 440 kv. Li resonance is <11 kv. wide. Experimental data are given by Bothe and Gentner (1937), Hafstad and Tuve (1935) and Curran *et al.* (1938), besides the references already cited.

The γ rays in our case were directly detected with Geiger counters (Craggs and Smee, 1941) shielded from the intense x radiation from the tube by about 2·5 cm. of lead. The absorption of γ rays in the shield is appreciable (Crane and Lauritsen, 1935). The Li radiation intensity is reduced by a factor of approximately 4 for 1 in. Pb. However, in view of the overwhelming amount of x radiation, a thick absorber is necessary. Few experimenters use <1 cm. Pb.

The targets used are described in table 3.

Table 3

Bombarded element	Target
Li	LiCl fused on to copper or molybdenum LiOH fused on to molybdenum Li evaporated *in vacuo* on to copper Li_2O vapour condensed on copper
F	CaF_2 fused on to molybdenum
B Be	Powdered and hammered into thin sheet lead
C	Aquadag baked on to copper

The stability of the counter was checked with a standard 0·2 mgm. radium source, 5 in. from the centre of the counter, before and after each run. Runs were taken with clean targets to assess the background count. The measured γ-ray intensities were never high, owing to the absorption in the counter shield and the small target current used. The C, B and Be yields were small compared with those from F and Li, and will not be discussed.

Representative runs are shown in figures 8 and 9. The focusing voltage was about 10 kv., to be added to the abscissae of the figures. Figure 9 was taken with F-contaminated LiCl. The voltage calibration of the apparatus was confirmed satisfactorily, but from time to time we noticed large deviations, due possibly to spurious charges on the corona screens affecting the generating voltmeter. We were not able to work with thin targets, because of the interruptions caused by more urgent work, but had developed the necessary evaporation technique.

Bothe and Gentner (1937) found relative yields from LiOH, CaF_2 and B of 30, 5, 1 at about 500 kv. Our figures for this voltage were 2·7, 16, 1. The possible errors were approximately ± 10 %, ± 10 % and ± 30 % respectively. We found the yields from CaF_2 to be consistently higher than those from LiOH, LiCl or Li_2O. The relative yields for Li, Li_2O, LiOH and LiCl were 8·5, 6·0, 3·6, 3·9. The discrepancies cannot be entirely due to experimental errors (about ± 10%). No correction has been made for absorption in the counter shield for either the Li or F radiations. The target solid-angles were maintained constant for all targets, and beam-size variations should be negligible. All yield figures given above refer to 500 kv. tube voltage.

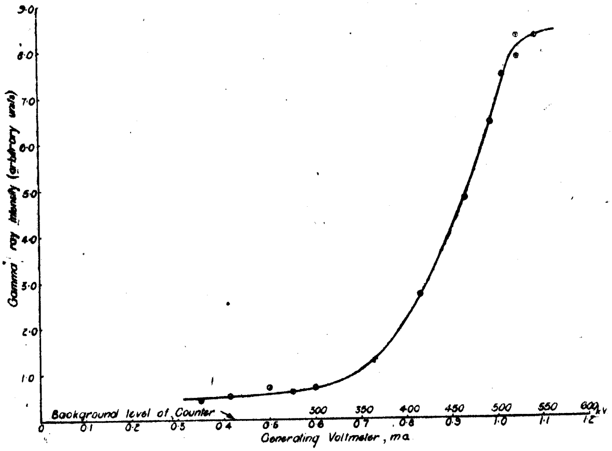
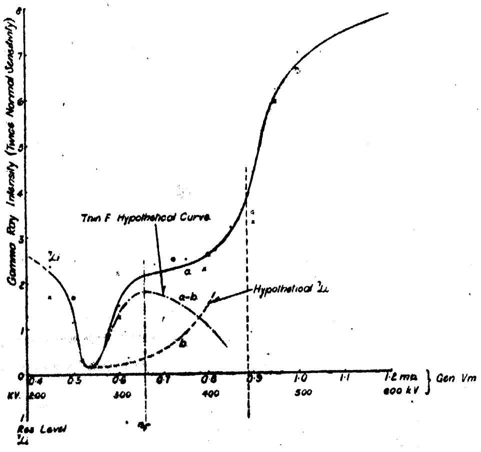


Figure 8. Evaporated lithium metal target.
30 μa. beam 0·2 Ra = 10·5 γ-ray units.

Fowler *et al.* (1938) give comparable data. They stated the ^{11}B radiation (180 kv.) to be much weaker than that from ^{19}F (330 kv.) or ^{7}Li (440 kv.). Corrected yields for LiOH and CaF$_2$ were respectively 7×10^{-9} and 5×10^{-9} quanta/proton. The CaF$_2$ yield was therefore higher than that found by Bothe and Gentner, using Geiger counters and 1·5 cm. Pb. shielding. The latter (1939) give the relative counter response for the Li and B rays as approximately unity.



Li Cl target (fused in molybdenum pan)
20 μa. beam
0·2 mgm. Ra ≏ 2 × 9·1 units

Figure 9. Gamma emission from fluorine-contaminated Li.

Hafstad and Tuve (1935) gave the Li-p output as equivalent approximately to 0·02 mgm. Ra/μa. at 500 kv. We performed an experiment with 0·2 mgm. Ra in a capsule (see above) fixed to the target, the 1-in. lead shield being removed from the counter, and obtained a count of about 4000/30 sec. LiCl bombarded with 35 μa. of mixed H_1^+ and H_2^+ ($\sim$50 % H_1^+) gave a count of 960/30 sec. at 550 kv., equivalent to 0·003 mgm. Ra/μa., i.e. approximately 0·007 mgm. Ra/μa. for an Li target. If we correct for the absorption of the Li γ rays in the counter shield, the agreement with Hafstad and Tuve data is reasonably good. More modern data for the F-p reaction (Streib *et al.*, 1941; van Allen and Smith, 1941) give the yield at 300 kv. as approximately $1{\cdot}4 \times 10^{-8}$ γ rays/proton, i.e. 9×10^4 γ rays/μa. sec., or about $1{\cdot}5 \times 10^{-3}$ mgm. Ra/μa. Unfortunately an accurate comparison between Li and F yields was not made by these authors.

In order to obtain more information on this matter we briefly investigated the absolute sensitivity of the Geiger counters, a knowledge of which might eliminate the direct calibration against radium. It is known that the γ radiation from light elements, in particular F (van Allen and Smith, 1941), is emitted equally in all directions from our target; hence a knowledge of the target solid-angle subtended at the counter, and the absolute sensitivity of the latter for the radiation considered, enable absolute yields to be determined. The method of Street and Woodward (1934) for calibrating cosmic-ray counters seems the most suitable for use with high-energy γ rays, but, due to lack of time, we were not able to perform similar experiments. The situation with regard to absolute efficiencies of γ-ray counters has been briefly outlined (Craggs, 1941), summarizing our own and other authors' results (e.g. Dunworth, 1940). For $\sim$2 Mev. γ rays the absolute efficiency of a normal counter is about 0·3 to 1 % and for our counters $\sim$0·5 %. The target solid-angle was estimated from counting experiments with radium (see above) to be about $0{\cdot}7 \times 10^{-2}\pi$. The count obtained with a CaF_2 target (5 μa. protons) at about 330 kv. was 520/20 sec. Assuming a counter efficiency of 1 per cent, the total count for a solid angle 4π was therefore $3 \times 10^5/\mu$a. sec., higher by a factor of 3 than the value of van Allen and Smith (1941). The error in assuming 1 % counter efficiency from radium measurements may be high enough to account for the discrepancy. As stated above, we were unable to carry out further experiments which, we hoped, would have given results of far greater accuracy.

§ 8. CONCLUSIONS

The development of a high-voltage apparatus for atomic disintegration work presents many additional problems when the available high-tension current is limited. These aspects of the technique have been specially emphasized, as it is believed they were not fully appreciated in other work.

Our experiments, unfortunately, are incomplete owing to the existing special circumstances, but target currents of 100 to 200 μa. have been obtained with 1·8 ma. high-tension generator current. This represents a good performance, despite the extremely low efficiency of the apparatus. When operated as a neutron generator, the set should give an activity of about 100 curies (Rn–Be), based on the data given by Heyn (1938): this is adequate for most purposes.

The original programme of work included certain experiments with neutrons—in particular, development work on neutron-sensitive Geiger counters and on various nuclear γ-ray effects.

§9. ACKNOWLEDGEMENTS

The author wishes to thank his colleague, J. F. Smee, for collaborating in some of the experiments described above. The work was carried out under the direction of Professor C. D. Ellis, F.R.S., to whom the author expresses gratitude for constant advice and encouragement. The London University Grants Committee provided means for purchasing the apparatus for ultimate use in King's College, London. However, many of the incidental expenses have been borne by the Metropolitan-Vickers Electrical Co., Ltd. Thanks should also be expressed to the Department of Scientific and Industrial Research for a grant which enabled part of the above work to be carried out; to Dr. T. E. Allibone, in whose laboratory the apparatus was developed, for his interest and suggestions; and to Dr. A. P. M. Fleming, Director of Research, Metropolitan-Vickers Electrical Co., Ltd., for permission to publish this paper.

REFERENCES

Allen, J. S., 1939. *Phys. Rev.* **55,** 336.
Bothe, W. and Gentner, W., 1939. *Z. Phys.* **92,** 45.
Bouwers, A., Heyn, F. A. and Kuntke, A., 1937. *Physics,* **4,** 153.
Craggs, J. D., 1941. *Nature, Lond.,* **148,** 661.
Craggs, J. D., 1942. *Proc. Phys. Soc.* **52,** 245.
Craggs, J. D. and Smee, J. F., 1942. *Proc. Phys. Soc.* **52,** 245.
Crane, H. R. and Lauritsen, C. C., 1935. *Int. Conf. Phys. Report* (*Phys. Soc.*), **1,** 130.
Curran, S. C., Dee, P. I. and Petržílka, V., 1938. *Proc. Roy. Soc.* A, **169,** 269.
Dunworth, J. V., 1940 *Rev. Sci. Instrum.* **11,** 167.
Fowler, W. A., Gaerttner, E. R. and Lauritsen, C. C., 1938. *Phys. Rev.* **53,** 628.
Goodlet, B. L., Edwards, F. S. and Perry, F. R., 1931. *J. Instn. Elect. Engrs,* **69,** 695.
Hafstad, L. R. and Tuve, M. A., 1935. *Phys. Rev.* **48,** 306.
Hailer, C., 1938. *Wiss. Veröff a.d. Siemens-Werken,* **17,** 115.
Herb, R. G., Parkinson, D. B. and Kerst, D. W., 1937. *Phys. Rev.* **51,** 75.
Hill, A. G., Buechner, W. W., Clark, J. S. and Fisk, J. B., 1939. *Phys. Rev.* **55,** 463.
Kaye, G. W. C., Bell, G. E., Binks, W. and Perry, W. E., 1940. *Rep. Progr. Phys.* **6,** 95.
Kirkpatrick, P., 1932. *Rev. Sci. Instrum.* **3,** 430.
Kossel, W., 1938. *Z. Phys.* **3,** 264.
Lauritsen, T., Lauritsen, C. C. and Fowler, W. A., 1941. *Phys. Rev.* **59,** 241.
Neubert, V., 1938. *Z. Phys.* **110,** 334.
Peek, F. W., 1929. *Dielectric Phenomena in High Voltage Engineering* (McGraw-Hill).
Street, J. C. and Woodward, R. H. 1934. *Phys. Rev.* **46,** 1029.
Streib, J. F., Fowler, W. A. and Lauritsen, C. C., 1941. *Phys. Rev.* **59,** 253.
Trump, J. G. and van de Graff, R. J., 1937. *J. Appl. Phys.* **8,** 602.
Van Allen, J. A. and Smith, N. A., 1941. *Phys. Rev.* **59,** 501.
Van Atta, L. C. and Nowak, W. B., 1941. *Rev. Sci. Instrum.* **12,** 305.
Van Atta, L. C., Northrup, D. L., Van Atta, C. M. and van de Graaff, R. J., 1936. *Phys. Rev.* **49,** 761.
Wintz, H. and Rump, W., 1931. *League of Nations (Health Organization).*
Zinn, W. H., 1937. *Phys. Rev.* **52,** 655.

A MACHINE FOR THE COMPUTATION OF STRUCTURE FACTORS

By R. C. EVANS, M.A., Ph.D., B.Sc.
AND
H. S. PEISER, B.A.,
Department of Mineralogy and Petrology, Cambridge

MS. received 30 *December* 1941 ; *in revised form* 15 *April* 1942

ABSTRACT. A machine is described which evaluates mechanically the expression $2f \cos(hx+ky)$, and its application to the computation of structure factors is discussed.

§ 1. INTRODUCTION

During the course of an x-ray analysis of a crystal structure, occasion repeatedly arises to calculate the structure factor of a proposed atomic arrangement, both when determining the signs to be attributed to the amplitude of each reflection, in order that Fourier syntheses may be prepared, and when confirming the proposed structure by an extensive comparison between observed and calculated intensities. Although the calculation of the structure factor is a very simple operation, it becomes tedious and time-consuming when repeated for a large number of reflections. The machine here described is designed to effect these calculations mechanically, and it has been found that under favourable conditions a very substantial saving in time is achieved by its use. In its present form, owing to lack of supplies of alternative materials, the machine is built almost exclusively of standard Meccano parts; it is, nevertheless, robustly constructed, and it has been found possible to realize a surprisingly high degree of accuracy, adequate for all the normal requirements of crystal-structure analysis.

The machine is primarily designed to derive the contribution F_{hk0} to the structure factor of the $(hk0)$ reflection due to two atoms, of atomic scattering factor f, situated at the points $\pm(x, y, -)$ and, therefore, related by an effective centre of symmetry at the origin $(0, 0, -)$. That is to say, the machine evaluates the expression

$$F_{hk0} = 2f \cos(hx+ky),$$

in which the co-ordinates x and y are expressed in degrees as fractions of the corresponding translations in the unit cell. The structure factor is required explicitly in this form in many analyses, but, as will be shown below, the machine can also be used to compute structure factors in the most general case. The structure factor of the structure as a whole, when required, must in all cases be obtained by summing the contributions computed by the machine for each pair of atoms.

§ 2. PRINCIPLES OF DESIGN

The mechanical principles of the machine are illustrated purely schematically in figure 1. The machine is operated by handles (A) and (B) which are turned through a number of revolutions proportional to h and k respectively. The multiplication of h and k by x and y is achieved by variable friction drives between discs (C) and (D) and wheels (E) and (F), the displacements of which from the centres of the discs are proportional to x and y respectively. The angular displacements of the shafts (G) and (H) are thus proportional to hx and ky respectively, and these displacements are added in a differential gearing (I), the drive from the cage of which is so arranged that the rotation of a calibrated drum (J) and a crank (K) is $(hx+ky)$ degrees. A carriage (L) is held against this crank and is constrained to move in one dimension: its displacement is therefore proportional to $\cos(hx+ky)$, the resolved part along its direction of motion of the rotation of the crank. Finally, the multiplication of $\cos(hx+ky)$ by the variable parameter $2f$ is achieved by a pointer (M), pivoted about a fixed point on the carriage and sliding through a fulcrum (N), the deflection of the end of which is observed on the scale (O). The amplitude of this deflection can be altered by displacing (N) in a direction perpendicular to the length of the scale, and in this way variations in $2f$ can be accommodated.

§ 3. DESCRIPTION

The machine is illustrated in figures 2, 3 and 4.

The values of h and k are introduced by two handles (1), the drives from which are transmitted to steel-faced plywood discs (5) whose rotation is restricted to integral numbers of complete revolutions by the location on heart-shaped cams (6) of rollers mounted on spring-loaded jockeys (22). The values of h and k up to $\pm$ 20 are recorded on graduated drums (23) which can be reset to zero without the necessity of rotating the handles (1).

The discs (5) are spring-loaded against brass wheels (2) having sharp rims and carried on shafts (4) co-planar with and perpendicular to the axes of the discs. These wheels can be locked to their shafts (4) in any position, the displacement from the centres of the discs (5) being determined by the parameters x and y, the values of which are read on the scales (3). Accurate settings are achieved by the screws (20), whose hardened conical points form the outer bearings of the shafts (4); by rotating these screws the shafts, and with them the wheels (2), are displaced through small distances.

By relieving the spring pressures of the friction drives by means of the handles (8), the shafts (4), and with them the rest of the mechanism, can be operated independently of the handles (1) by turning the wheels (21).

The angular displacements of the shafts (4) are added together in the differential gearing (28), the rotation of which is transferred to a graduated drum (7) and to a crank (18) whose angular displacements are arranged to be $(hx+ky)$ degrees.

The carriage carries on its underside three parallel rails (27, 37 and 35) which rest on three wheels (32, 36 and 33) of a lower carriage. These wheels in their turn rest on fixed rails (31, 39 and 19), and the displacement of the carriage is geometrically constrained to motion in a straight line. A bar (34) on the carriage is pressed against a roller (30) on the crank (18) by the action of a spring (29), and the displacement of the carriage is therefore proportional to $\cos(hx+ky)$.

A pointer (10), pivoted at its centre of gravity on the shaft (17), is attached to the carriage and slides through a block (16) which acts as its fulcrum. At its

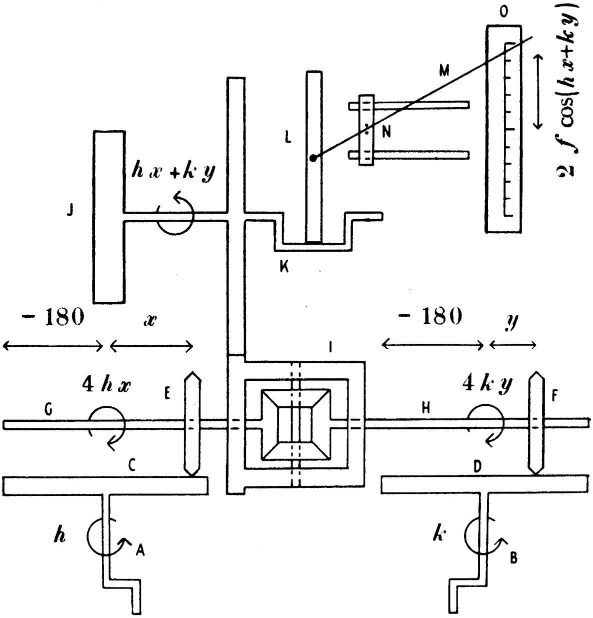


Figure 1.

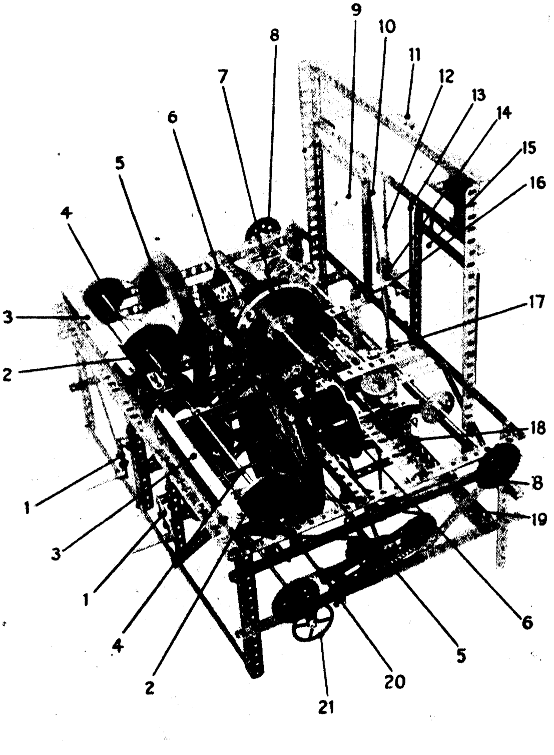


Figure 2.

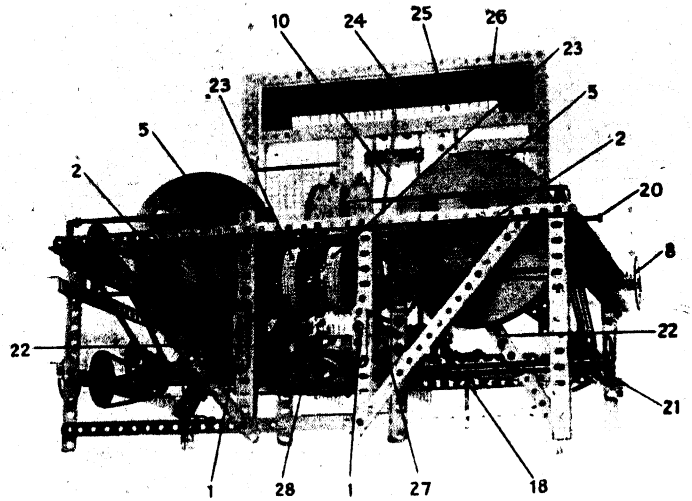


Figure 3.

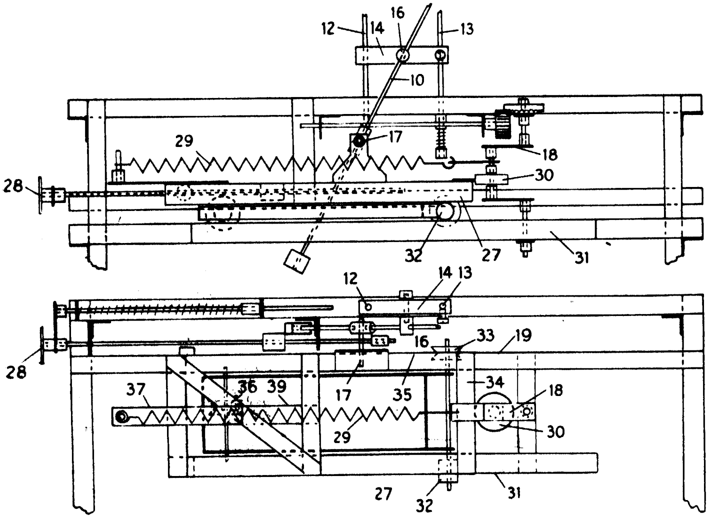


Figure 4.

upper end the pointer carries a thin needle (24) moving over a scale (25) fixed in front of a mirror (26) to enable readings free from parallax to be taken. The height of the fulcrum (16) above the pivot (17) can be adjusted by sliding the block (14) on two vertical guides (12 and 13). A fine adjustment of this setting is provided by a nut (11), and is facilitated by a handle (28) which displaces the carriage into the position corresponding to $\cos(hx+ky)=1$, so that the value of $2f$ may be read directly on the scale (25) without disturbing the handles (1).

Back-lash has been practically eliminated by spring-loaded gear wheels and by means of suitable jockeys for the chain drives. Care has been taken to reduce friction and inertia, especially beyond the friction drive, all danger of slip in which is eliminated by a 2 : 1 step-down between the handles (1) and the discs (5). The force required to drive the machine when applied to the rims of the wheels (2) is about 30 gm. weight.

§ 4. MANIPULATION

The values of x, y and $2f$ are introduced by the adjustment of the wheels (2) and the fulcrum (16). The scale (7) and the drums (23) are set to zero and the machine is ready for use.

The structure factor F_{hk0} for the $(hk0)$ reflection is displayed on the scale (25) after turning the handles (1) to the appropriate values of h and k; these values are recorded on the drums (23). Unnecessary effort in taking the readings may be avoided if the planes are arranged in a convenient sequence. The value of $(hx+ky)$ may be checked regularly, and if necessary corrected, by a direct comparison between the values of $(hx+ky)$ recorded on the scale (7) for the (500), (10, 00), (050), (0, 10. 0), etc., reflections and those given in the tables (9 and 15) which record the values of $5x$, $10x$, $15x$ and $20x$. In practice it is rarely found necessary to make any such correction unless the calculations are extended over several hundred reflections involving, in some cases, a total displacement of the scale (7) of some tens of thousands of degrees.

In many cases it will be unnecessary to make any allowance for changes of f with $(\sin\theta)/\lambda$. If, however, it is desired to make such an allowance, the computation may be interrupted at any stage and the value of $2f$ readjusted. In such cases it may be convenient to carry out the calculation in such a sequence that those reflections for which a single f value can be used are grouped together, since the time consumed in turning the handles (1) through a few revolutions is small compared with that required to reset $2f$.

In cases in which it is required to evaluate the structure factor F_{hkl} of the general plane (hkl), given by

$$F_{hkl} = 2f\cos(hx+ky+lz),$$

the calculations may be carried out as described above for the groups of planes $(hk0), \ldots (hkl), \ldots (hkL)$ successively. For this purpose it is only necessary to pre-set the scale (7) to the value of lz appropriate to each such group instead of to zero. If the number of reflections is large, the time spent in making these adjustments is quite insignificant. Values of $F_{hk\bar{l}}$ may be deduced from the relation $F_{hk\bar{l}} = F_{\bar{h}\bar{k}l}$ to avoid the necessity of setting the scale (7) to both lz and $-lz$.

Although the machine lends itself most readily to the computation of the structure-factor contribution due to two atoms related by a centre of symmetry, it can also be used in the general case of the space group 1. In this case the structure factor of a single atom is equal to f, but for several atoms the structure factor is given by

$$\begin{aligned} F^2_{hkl} &= \{\Sigma f \cos(hx+ky+lz)\}^2 + \{\Sigma f \sin(hx+ky+lz)\}^2 \\ &= \{\Sigma f \cos(hx+ky+lz)\}^2 + \{\Sigma f \cos(hx+ky+lz-90^\circ)\}^2, \end{aligned}$$

where the summation is carried out over all the atoms in the structure. In this expression each of the two terms in brackets may be evaluated by considering each atom in turn. For the latter term it is only necessary to set the scale (7) initially to 270° instead of to zero and then to proceed as before. The same preliminary setting of the parameters x, y and z will, of course, serve for both terms of the equation.

§ 5. EVALUATION OF STRUCTURE-FACTOR CONTRIBUTIONS OF ATOMS RELATED BY SYMMETRY OPERATIONS OTHER THAN A CENTRE

In those non-centrosymmetrical and centrosymmetrical space groups in which the numbers of equivalent general positions exceed one and two respectively, the machine can be employed to compute the structure-factor contribution due to each atom or pair of atoms separately. In general, however, much of the labour that would be involved in such calculations may be avoided. Thus, for example, in the space group $P2/m$, in which the structure factor is given by

$$F_{hkl} = 4f \cos(ky) \cos(hx+lz),$$

the calculations may be carried out successively for groups of planes with a common index k. For each such group, $4f\cos(ky)$ may be introduced as a constant quantity by the adjustment of the fulcrum (16), and if the number of planes for each value of k is large the saving in time achieved by the use of the machine is as great as in the simplest case of two atoms related by a centre. If however, the structure factors of planes of the type ($hk0$) and ($0kl$) alone are required, the advantage of using the machine is substantially less. Closely analogous arguments are applicable to all other space groups of the three classes of the monoclinic system, and will immediately suggest themselves on consideration of the relevant expression for the structure factor.

In many of the space groups of the orthorhombic system, the structure factor is in a form which lends itself less immediately to calculation by the machine. In all these space groups, however, there is, of course, at least one diad axis, and it is therefore always possible to apply to groups of atoms related by this axis arguments similar to those employed in the monoclinic system. In space groups of systems of higher symmetry, in which the number of equivalent positions is greater, the advantages of the machine are correspondingly reduced; in almost all cases, however, it may conveniently be employed to perform at least part of the calculations involved.

§ 6. RESULTS

Some typical results obtained by using the machine are recorded in table 1. For these observations the structure factors $9{\cdot}3\cos(h.154-k.78)$ and $9{\cdot}3\cos(h.32+k.172)$, corresponding to the arbitrarily chosen parameters (154°, −78°, −) and (32°, 172°, −), were computed for the sequence of 154 planes (000), (100) . . . (15, 00), (15, 10) . . . ($\overline{15}$, 10), ($\overline{15}$, 20) . . . (15, 20), (15, 30) . . . (15, 50) . . . (050), and the values obtained for the first and last and for two intermediate groups of five successive planes in this sequence are recorded. The calculated values of the structure factors, together with the observed and calculated values of $(hx+ky)$, are also shown in the table. From the figures it will be seen that the maximum error in F_{hk0} is 0·5—for (250)—and that this error is associated, as is to be expected, with a small value of F. The value of $(hx+ky)$ is in no case in error by more than 2°. An analysis of the complete set of 308 observations shows the errors in F to be distributed as follows:—

Table 1. Accuracy obtainable in a typical calculation

$hk0$	$h.154-k.78$		F_{hk0}		$h.32+k.172$		F_{hk0}	
	Obs.*	Calc.†	Obs.	Calc.	Obs.	Calc.	Obs.	Calc.
	°	°			°	°		
100	154	154	−8·3	−8·4	32	32	7·8	7·8
200	307	308	5·6	5·7	64	64	4·2	4·1
300	100	102	−1·7	−1·9	96	96	−0·8	−1·0
400	254	256	−2·6	−2·2	128	128	−5·6	−5·7
500	50	50	6·0	6·0	160	160	−8·9	−8·7
$\bar{5}20$	154	154	−8·3	−8·4	184	184	−9·3	−9·3
$\bar{4}20$	308	308	5·7	5·7	215	216	−7·7	−7·5
$\bar{3}20$	102	102	−1·9	−1·9	248	248	−3·7	−3·5
$\bar{2}20$	254	256	−2·6	−2·3	279	280	1·5	1·6
$\bar{1}20$	50	50	5·9	6·0	312	312	6·3	6·2
$\bar{5}40$	356	358	9·3	9·3	168	168	−9·1	−9·1
$\bar{4}40$	150	152	−8·0	−8·2	200	200	−8·8	−8·8
$\bar{3}30$	304	306	5·2	5·5	231	232	−5·9	−5·7
$\bar{2}40$	99	100	−1·3	−1·6	263	264	−1·0	−1·0
$\bar{1}40$	252	254	−3·0	−2·6	296	296	4·1	4·1
450	224	226	−6·8	−6·5	268	268	−0·4	−0·3
350	70	72	3·2	2·9	236	236	−5·2	−5·2
250	276	278	0·8	1·3	204	204	−8·6	−8·5
150	122	124	−4·9	−5·2	172	172	−9·2	−9·2
050	329	330	8·0	8·1	140	140	−7·1	−7·1

* i.e. obtained by means of the machine.

† i.e. calculated from $2f\cos(hx+ky)$.

Error	0·5	0·4	0·3	0·2	0·1	0·0
Frequency of occurrence . . .	2	14	24	63	105	100

The accuracy of the machine is thus amply adequate for all normal requirements of structure analysis.

The time saved by the use of the machine was estimated by computing the structure factor of two groups of planes, containing about 300 and 600 reflections respectively, both with the aid of the machine and by numerical methods. For the latter calculations, standard tables of hx and ky, reduced to angles in the range 0 to 360°, were employed to give $(hx+ky)$. The value of $\cos(hx+ky)$ was taken directly from trigonometrical tables and the multiplication by f was carried out on two slide rules. The results of this comparison are shown in table 2. It is seen from these figures that the time saved by the machine is about 60% of that required to make the calculations from tables. Where, however, it is necessary to take account of variation of f with $(\sin\theta)/\lambda$ the advantage of using the machine is substantially reduced. On the other hand, much of the time spent in performing calculations with the machine is employed in recording the observations, and if an unskilled assistant is available to transcribe the results as they are dictated, a further substantial saving in time is achieved. Thus it was found that 600 observations could be recorded in this way in 41 minutes, compared with 86 minutes when the operator worked alone. A further 8 minutes was required for setting the x and y values, so that the total time for the observations was 49 minutes, compared with 240 minutes when tables were used.

Table 2. Speed attained in a typical calculation

Machine			Numerical methods		
	300 planes	600 planes		300 planes	600 planes
Setting x and y	10 min.	8 min.	Computing $(hk+ky)$	31 min.	70 min.
Recording F_{hk0}	31 ,,	86 ,,	Evaluating F_{hk0}	77 ,,	170 ,,
Total	41 ,,	94 ,,	Total	108 ,,	240 ,,

In conclusion it may be pointed out that computations can be performed with less fatigue with the help of the machine than by numerical methods. The chance of accidental error is, moreover, reduced by its use.

§ 7. ACKNOWLEDGEMENTS

It is a great pleasure to express our thanks to Meccano, Ltd., and especially to Mr. F. Riley, B.Sc., who made available to us the parts used in the machine, and to Mr. A. N. Lanham, who constructed those parts which could not conveniently be made of Meccano. One of us (H. S. P.) is indebted to the Council of St. John's College, Cambridge, for a Hutchinson Studentship.

REVIEWS OF BOOKS

The Nature of Thermodynamics, by P. W. BRIDGMAN. Pp. xii + 229. (Cambridge Mass.: Harvard University Press; London: Sir Humphrey Milford at the Oxford University Press, 1942.) $3.50. 20*s*. net.

In this striking book, Professor Bridgman analyses the concepts of thermodynamics from the " operational " point of view which he has described in previous writings and which maintains that the meaning of a physical concept is wholly expressed by the set of operations by which the concept is determined. Though there are still wide divergences of opinion on the matter, it is probable that the great majority of physicists have been much influenced by the operational outlook, and they will accordingly turn with interest to this minute analysis of the laws of thermodynamics which, as the author truly says, smell more of their human origin than other physical laws, and have also the distinction of being the only ones which have retained both their significance and their importance without appreciable change since the mid-nineteenth century.

It may be said at once that probably no such searching analysis of thermodynamic concepts has previously been made, and that anyone who can peruse this book carefully without clarification of his own ideas is either incredibly sagacious or essentially ineducable. It is the kind of book that should be read first fairly rapidly and then again, at least once, with more care. If the following remarks are almost entirely critical, that is because such treatment is the most fitting tribute that can be paid to a most stimulating attempt to bring order into the most elusive group of physical ideas.

The bulk of the discussion is contained in two chapters dealing with the first and second laws respectively ; two shorter chapters are given to miscellaneous questions. Throughout the book, Bridgman rightly insists that macroscopic are more fundamental than microscopic considerations, since all instrumental operations are necessarily macroscopic ; hence it is invalid to reject any conclusion properly drawn from thermodynamical laws simply because it does not fit into the present form of statistical or kinetic theory. This needs emphasizing, but it is unfortunate that—doubtless through an oversight—after the remark that " the task of thermodynamics is to give us mastery of what happens . . . . when temperature differences appear in the system ", the essence of thermodynamics is introduced in such an anticlimax as " A thesis of thermodynamics is that certain very broad aspects of the new and extended phenomena can be exhaustively handled by the introduction of the temperature of the various parts of the system ". It should be added that the succeeding analysis has infinitely more substance than this would lead us to expect.

The first law is expressed in the form, $\Delta E = \Delta W + \Delta Q$, and attention is concentrated on the operational meaning of the three terms in this equation. Bridgman recognizes that the actual use of the law goes beyond what can be strictly justified operationally, but he is less concerned to condemn this than to understand exactly what the actual use implies. He finds in it much " verbalising "—i. e. not merely " pencil and paper " operations, but also elements of the mental pictures which we associate with the concepts, not all of which have operational sanction. He views this, and also an occasional " mildly vicious circularity " of argument, with a leniency which will command much sympathy from physicists more interested in acquiring new territory than in establishing their right to it ; but, as events have shown more than once in the history of physics, logic is a jealous goddess, and indifference to her demands is liable to meet with retribution.

For example, in an interesting discussion of the localization of the energy " transmitted " from a motor through a belt to a flywheel, Bridgman introduces a Poynting vector to represent the mechanical energy flux, and later finds himself compelled to " say " (he admits that he is verbalizing) that the energy flux must be along the belt. " If instruments outside the belt give no readings ", he says, " then I have to *say* that the path of energy flow is in the belt ". It is a perilous argument. He later gives reasons for believing that there should be a field of force (obviously extremely weak) related to a gravitational field in the same way as a magnetic is to an electric field, and so a gravitational Poynting energy vector. If such a field should accompany the stress in the belt, the fact that external instruments give no readings would be due merely to their insensitiveness, and the flow of energy might be directed, as with a wire carrying an electric current, from the surrounding medium into the sides of the belt. The point is that the *direction* of energy flow is independent of the *magnitude* of the energy components, and a verbalization which relates them invites disaster.

The whole treatment of the localization of energy, which forms the backbone of the discussion of the first law, seems to me to be based on an error. Modifying a view taken in an earlier book (*The Logic of Modern Physics*), Bridgman feels compelled to ascribe " full and complete " localization to energy because, he holds, the first law is applicable to any region, whether occupied by matter or not, defined by a surface which separates it from its surroundings. By drawing the surface in all possible places we are therefore forced to grant complete localization to energy. Here it seems that he has been betrayed into forgetting his own operationalism. He is at infinite pains to elucidate the operational procedure for measuring the ΔW and ΔQ that cross the boundary (even so, it is not easy to see how to apply the method of measuring ΔQ in empty space), but he nowhere touches on the impossibility of operationally identifying the boundary itself. That this is impossible in space unoccupied by matter is the plain meaning of the special theory of relativity, and if you cannot localize a surface, then clearly you cannot localize the energy inside it. Hence the whole argument for energy localization collapses.

Indeed, one is inclined to think that the greatest obstacle to Bridgman's quest for understanding is his barely disguised dislike of the viewpoint of relativity, notwithstanding that it was the confessed origin of his own operational method. His references to relativity constitute a sort of gesture of disapproval rather than an open attack, and he often loses the significance of the relativity outlook by regarding its requirements as merely a " small correction " which may be ignored. This is a fatal attitude to the problem of understanding physical concepts, reminiscent of that of the erring damsel who excused her misdemeanour on the ground that the baby was only a very little one. As a prominent example we may cite the remark, on which much is based, that " the mass of a stone is unchanged (except for relativity effects) when it is measured in a frame of reference moving with respect to the original frame, but its kinetic energy is drastically altered ". The fact is that the changes of mass and kinetic energy are precisely equivalent, but are represented by numbers of different magnitude simply because different units of measurement are chosen. One gram is c^2 ergs, and to say that the change of mass is negligible compared with the change of kinetic energy is tantamount to saying that the change in E arising from ΔQ is negligible compared with that arising from ΔW because there are 40 million ergs in a calorie. Incidentally, Bridgman nowhere mentions that the first law, in the form given, is true only when the units of measurement of ΔW and ΔQ are the same. As measured by the operational processes described, they would normally be obtained in very different units, and it is at least arguable that the essence of the first law lies in the fact that ΔW can be added to ΔQ.

The discussion of the second law is largely parallel to that of the first, with entropy occupying the place previously taken by energy. Thus the sense in which we can speak of a localization and a flow of entropy is examined, and limits are sought to the legitimacy of ascribing " physical reality " to entropy. There is a very suggestive discussion of the meaning of reversibility and recoverability, and the statistical aspect of entropy is considered at some length. A possible extension of the second law to cover irreversible phenomena is adumbrated ; according to this, the irreversibility of a process would be represented by a definite assignable increase of entropy, instead of an indefinite one as in the ordinary formulation of the law.

There is much in this chapter on which one would like to comment. One point must suffice, namely, the conclusion that black-body radiation accompanied by fall of temperature is reversible. This is certainly a startling result, which if true, is of fundamental importance. We give the argument in Bridgman's own words :—

> " Idealise the universe as a sphere at uniform temperature which has been prevented from radiating since its creation by an impervious shield. Remove this shield, allowing the sphere to radiate a finite ' gulp ' of radiation, and replace the

> shield. The entropy of the sphere has decreased, but to compensate there is entropy traveling in the gulp of radiation. Next catch this gulp of radiation on a perfectly reflecting spherical mirror placed at the confines of the 'universe', reversing the direction of radiation. The radiation eventually returns to the sphere, is there absorbed, and the sphere in the end resumes its initial temperature. The initial entropy of the sphere is therefore also resumed. There is no other place where there may be any entropy, so that the net result of the emission and re-absorption of the gulp must be that the total entropy of the universe is unchanged. But since the entropy of a closed system can never decrease, this means that the total entropy was always the same. The entropy associated with the gulp of radiation while it was traveling through space was exactly equal to the entropy lost by the sphere. Furthermore, since the occurrence of an irreversible process leaves its indelible impress in a permanent increase of entropy, it must be that no irreversible or entropy-increasing processes occurred. That is, neither the emission of radiation by the sphere in the first place nor its re-absorption later is to be described as an irreversible process. This, I believe, is quite different from what would be the snap judgment of many physicists. It is evidently the unidirectional character of the radiation that distinguishes this special situation. Although the unidirectional emission of radiation is a process which occurs spontaneously and at a finite rate it is not an irreversible or entropy-increasing process, as is the conduction of heat down a temperature gradient, a process which to our first crude physical intuition might 'feel' to be the same."

Bridgman goes on to say :—

> "The entropy increase arising from .... exchange of radiation between bodies at different temperatures is not to be sought in the initial act of absorption, which may be non-entropy-increasing, but is to be found after the initial absorption in the spreading out of the spectrum of the absorbed energy from the distribution characteristic of the higher temperature of its source to the distribution characteristic of the lower temperature of the sink."

This argument does not appear to me to be sound. In the first place, the radiation is not "unidirectional"; it proceeds in all directions from the source. "Symmetrical" or "isotropic" is the word required. This is not merely a verbal difference, because whereas unidirectional radiation is unusual, and would justify the phrase "special situation", radiation outwards from the source is normal, and in this respect the artificial situation envisaged is in no way "special". But the chief criticism is that the process of removing the screen to re-admit the returning radiation would necessarily allow another "gulp" to escape, and the original situation would not be restored. And, in fact, in the nature of things it cannot be restored, for there is no conceivable way of restoring it without the use of a "shield" which transmits all radiation in one direction, but none in the opposite one. We have no experience of which this can be regarded as a legitimate extrapolation. We may admit a perfect reflector since we can realize it with close approximation, but a one-way transmitter is a sheer fantasy. Consequently, when the first gulp of radiation left the universe something physically irreversible happened, and the entropy must accordingly have increased.

There is another aspect of the passage quoted which is of considerable significance, namely, the ascription of the entropy increase when radiation is absorbed to the re-adjustment of the spectrum of the absorbed radiation, and not to the act of absorption itself. This shows a considerable departure from the "operationalism" of *The Logic of Modern Physics*, for how can one distinguish operationally between the initial absorption of radiation and the re-arrangement of its (obviously unobservable) spectrum while it remains absorbed? The distinction is purely verbal, and it is with surprise and some regret that we find Bridgman not only adding "paper and pencil" to "instrumental" operations, but also including the formation of a mental picture among the operations which give significance to a physical concept. Thus on pages 163–4 he writes :—

> "One might be tempted to say that 'operationally', since there is nothing to distinguish between the two views, there is in fact no difference, and that they must be regarded as the same. This would indeed be true, only 'operationally'

> must cover *all* the operations, not merely those of numerical calculation and program drawing. In particular, the operations include the entire verbal structure into which the two views may be fitted. From this point of view there is a difference which must not be overlooked between the common-sense point of view which says that a pail of water has never frozen on the fire, and I shall make my programs for the future without leaving any room for this actually happening although I recognize it as a *possibility* that it *might* happen, and the point of view that says that there is a chance of 10^{-1000} that a pail of water might freeze on the fire but this chance is so small that it would be waste of time to figure what course of action I had better adopt if it did happen ".

One cannot help feeling that in thus enlarging the meaning of " operation ", Bridgman is losing the chief value of the operational point of view. As originally formulated, that point of view did make clear a sharp distinction between possible directions of physical speculation, but if operations are to include the " mythology " with which we find it comforting to gild the philosophic pill, it is difficult to see that any distinction at all is left. Even absolute simultaneity passes the operational test by virtue of its unique " verbal structure ". One may grant that an insistence on instrumental operations alone would prohibit progress along directions that might prove both legitimate and profitable. " Paper and pencil " operations, in the sense of strictly specified mathematical procedures, would seem to be necessary for the full development of physics. But Bridgman's pre-occupation with " physical reality ", which he admits surprises himself, indicates a shift of interest from the conceptions of physics to the misconceptions of physicists. That is a misfortune which it may be hoped is only temporary.

The possibility must be faced, however, that this *reductio ad absurdum*, as we might almost call it, is the inevitable conclusion of the rigorous development of the operational method, just as the logical positivists' original criterion of meaning has issued in realization of the nonsensical character of their own statements. It has always seemed to me that " operationalism " is a good servant but a bad master. It is excellent, for example, to realize that a measurement of a velocity in terms of the space covered in a given time is something different from a measurement in terms of the displacement of a spectrum line, and there are occasions (possibly the phenomenon of the " expanding universe " is an example) when recognition of the difference may be of cardinal importance. At the same time, the construction of a theory according to which the measurements are regarded as equivalent is a great step forward and is responsible for a great part of our present knowledge of astrophysics. A fundamental definition of physics demands, first of all, recognition of the distinction between the experiences which are our data and the rational concepts which we use to correlate them, and a doctrine such as operationalism which can be expressed only with difficulty in terms of this distinction is thereby marked as of doubtful fundamental validity. That is not to say that it is not of extreme value on its own level, and we end as we began by an acknowledgment that Professor Bridgman's study is one of the most valuable contributions to the understanding of thermodynamics that have yet appeared.

HERBERT DINGLE.

Science and Education (*Current Problems*, 15), by S. R. HUMBY and E. J. F. JAMES. Pp. viii + 146. (Cambridge University Press, 1942.) 3*s*. 6*d*.

It has become increasingly evident that there is a great gap between what science has done and what it might have done for society. This book asserts that the gap is due to the unawareness by society of the social changes produced by science in the past and, therefore, of the possibility of the solution of our present social problems by science. It traces this unawareness partly to the absence of science from the curriculum of the classical sides of many public and grammar schools and partly to the unsatisfactory nature of the little science which is taught to the other boys in these schools. It therefore makes radical constructive suggestions about the content of, and method of presentation in, the science courses in schools and universities.

The book commences with an outline of the social changes of the last three centuries due to science, citing examples such as the growth of urban populations, the new structure of society produced by the industrial revolution, and the possibility of the age of plenty.

It suggests that the evils associated with great cities (their smoke, their slums, their disease) and the continued existence of poverty in the face of the possibility of plenty can be eradicated by the application of scientific knowledge and scientific method. The failure to do so is attributed both to the fact that most of our cabinet ministers and civil servants learned little or no science at school, and also to an unawareness on the part of society generally of what science has done and therefore could do to influence social change. While the evidence for the gap between what science has done and what it might have done for society is good, the arguments that this gap is due to unawareness are less convincing. For example, the authors state that " this fatal division between resources and unawareness is shown in any number of ways in almost every department of our life ". The reader expects a selection of striking examples, only to find an apocryphal story about Gladstone and Faraday! It may be that there is good evidence ; if so, would it not have been wiser to quote it ? And is the unawareness quite so glaring as the book suggests in view of the fact, quoted in the book, that science is now a profession, whereas a hundred years ago it was the pursuit of a few wealthy amateurs ? And does not the creation during the last fifty years of the National Physical Laboratory, the Department of Scientific and Industrial Research, the research departments of the various trade associations and of many private firms show that society is not altogether unaware of the value of science to industry, at least ?

Even if the explanation of the gap between the achievements and possibilities of science is not accepted, the suggestions made in the book about the modification of the content and presentation of science in schools and universities would help to reduce that gap and so deserve the most careful examination. The authors rightly condemn the paucity of time still devoted to science in public and grammar schools, and the prejudice which still exists against it ; this results both in a large proportion of the able boys being forced to do classics, and in a divorce between humane and scientific studies, the one being regarded as cultural and the other as merely utilitarian. The authors plead for a broader, more humane course of science for the future citizen (as distinct from the future scientist), which should include not only chemistry and physics, but also some biology, astronomy and geology, and should be linked with history by the study of the social changes produced by science. They emphasize the need to infuse the essentials of the scientific method into the future citizen, so that he may be able to distinguish truth from propaganda and be better able to weigh objectively and dispassionately the issues which a citizen must continually face. Finally, they recommend a course of general science of two periods a week after school certificate for *all* specialists, in which recent scientific knowledge and the application of science to modern social problems is discussed. They claim that if a future civil servant or cabinet minister had learned science in this way, he would know what type of social and political problem was capable of solution on scientific lines, and how much progress could be expected in a limited time ; he would, in fact, be able to judge when to call the scientist to his assistance.

The book is an interesting and critical discussion of a problem of vital and universal interest, and it should be read by university teachers and schoolmasters and should be in all university and school science libraries. A. W. B.

A History of Science, by Sir WILLIAM CECIL DAMPIER. Third Edition. Pp. xxiii + 574. (Cambridge University Press, 1942.) 25*s*. net.

The preface assures us that this is, in effect, a new book. It is at the same time a third edition and has the advantage which comes from revision and enlargement. But phrases like " in this year 1929 " do occur here and there, and suggest that it has not lost all traceable likeness to the second edition (1930).

The decade 1930 to 1940 is covered by the addition of an eleventh chapter, and as the work begins with the scientific thought of the ancient world it can be easily imagined what a stupendous task has been undertaken. But this in itself gives little idea of what has been attempted, for to the history of science is joined the development of philosophic ideas, which bulk quite largely ; and, as if this were not enough, considerable space is devoted to the discussion of religion. This last subject seems to conclude (p. 489) in the theory that ritual transcends dogma in importance. It is, perhaps, not altogether unfair to infer that the work is written from the angle of the college chapel, preferably with a good choir.

As it is quite impossible to cope at all adequately with an extremely interesting and suggestive work of such vast extent in a short review, it seems better to quote the author's own candid admission (p. 493), which can hardly be final: " Later views have brought into prominence older discoveries which at that date (1930) appeared of less interest. Hence, at some points, it is necessary to turn back in time and describe older investigations, ignored in the first two editions of this book ". But a little thought may perhaps be given to a question which arises naturally in a work of this kind, the division of time into significant historical epochs.

To approach this question in a concrete form, how should the eighteenth century be considered? The choice is made in the belief that this period has been most commonly neglected. Now, if there is a period of scientific history which is clearly defined, it is surely the seventeenth century, the age of Galileo, Huygens and Newton. Moreover, as the age of Des Cartes, Spinoza and Leibniz, it is equally defined from the philosophic angle. Thus, for want of a better, the year 1727 may be suggested as the beginning of the next epoch. How are the following years to be regarded? From the English point of view there is a strong temptation to pass them over lightly. The English universities, with their collegiate structure, suffered from the ills to which their constitution is peculiarly prone, and cut a very sorry figure. But the eighteenth century in science cannot be dismissed lightly. It is the age of Euler and Lagrange, it carries electricity from the terrestrial magnetism of Halley to the days of Faraday, and science, with the addition of technology, is widened in its entire scope. All these things bring another question. If, as it appears, the eighteenth century represents something which deserves to be considered on its merits, what limit in time should be imposed? A Cambridge man often likes to think of a *risorgimento* occurring about the year 1820. But this parochial view may be corrected by trying to see the matter through French eyes. In spite of revolution in the political sphere, no discontinuity is visible in science. In mathematics the progress from Clairaut to Lagrange is maintained from Cauchy to Poincaré without any palpable break, and elsewhere a similar continuity can be traced. Surely a clue is beginning to emerge in a single age of the classical mechanics and the unbroken atom, starting from the death of Newton, embracing the use of technology and ending about 1895. If it must be separated into two halves, the Victorian age may serve for the second, but there is no organic interruption. Should it be objected that this means giving too much weight to the influence of the physical sciences, the answer is that the others must either keep step or the attempt to treat the whole in one synthesis is doomed to failure. As time passes, it must become more difficult to preserve an unbroken front.

This scheme, it must be confessed, fares ill at the hands of Sir William Dampier, and the divergence begins early. It flows from the treatment of the seventeenth century, which is divided at 1660. The result is to isolate the Newtonian epoch (chap. iv) from all that went before and clear the decks for the nineteenth century, which is discussed in three separate chapters (v–vii) followed by four more (viii–xi) on still more recent times. On the other side, after the second chapter has traced the Middle Ages from the decline of Rome to the end of the thirteenth century, the interval from scholasticism at its height, with science feeling its way step by step, is continued past Galileo to the foundation of the Royal Society. Such arrangements are doubtless a matter of free choice, and the individual ought not to be hampered. But, according to the test which has been suggested, the treatment accorded to the eighteenth century appears rather less than its due.

Scientific opinion in general has little sympathy with philosophical ideas. Like the curvature of an arc, philosophy can be ignored in science so long as the range is sufficiently small. In the distant past it intrudes and must be taken into account. It occupies considerable space in this book, and this is natural enough. That philosophy is once more claiming attention in the science of the day is, perhaps, merely a symptom of adjustment of ideas to that curvature which may have been too long ignored and resisted. But the philosophy of the future is likely to differ widely from that of the past because, to satisfy the needs of science, it must be evolved from within, not imposed from without. Like mathematics, where the point can no longer be contested, science (if worthy of the name) must be master in its own house.

From this point of view the eighteenth century refuses to be treated as a part of the Newtonian epoch. For it is in this age that the claim of science to independence from philosophy and religion is not only made but conceded, however grudgingly, for the first time. The limited supremacy of Newton can also be illustrated by considering two individual astronomers, both of the first magnitude. James Bradley owes this eminence not to the discovery of aberration (the only achievement mentioned in this book), but to the fact that he belongs to that small class who have raised the general standard of accuracy in observation. He is much more closely allied to Flamsteed than to Newton. The other astronomer is Sir William Herschel, only mentioned here on account of an isolated experiment on the infra-red solar spectrum. In spite of his discovery of Uranus, his greatness lies in the fact that he looked far beyond the solar system, which was Newton's main concern, and enlarged the boundaries of the visible universe. Both these astronomers, in different ways, illustrate another characteristic of the age in which they lived, the growing importance of finer and larger instruments. But more than enough has been said on behalf of the eighteenth century to show that, even apart from the biological sciences, it cannot be fairly treated as a Newtonian preserve. H. C. P.

Introduction to Physical Chemistry, by ALEXANDER FINDLAY. 2nd Edition. Pp. vi+582. (London: Longmans, Green and Co., 1942.) 15*s*.

This second edition of a widely used book is some ninety pages longer than its predecessor. The new material is distributed fairly generally throughout: notice should be made of two new short sections on *Elements of Crystallography* and *Molecular Dimensions and Configurations*. The number of problems has also been increased. These additions result in some shifting of emphasis, and much of the discussion has been re-written and references to the literature have been brought up to date.

The presentation of any work of this kind is bound to reflect, in some way, the scientific pre-occupations of the author. It is not surprising, therefore, to find the chapter on the Phase Rule one of the best in the book.

As befits an introductory text-book, the discussion is general and descriptive, but it is, perhaps, a pity that there is so little of more precise treatment. Both kinetic theory and thermodynamics suffer somewhat in consequence. It is also unfortunate that (p. 293) a distinction is made between thermodynamics and thermochemistry, whereby the former has the now standard " acquisitive " sign-convention, the opposite convention being used for " thermochemistry ". Surely the former convention only is required.

The first edition provided a well-written introduction to the subject: the matter added to the new edition has brought it into line with recent developments, and increased its usefulness. R. F. B.

Biography of the Earth: Its Past, Present and Future, by George Gamow. Pp. xiv+242. (New York: The Viking Press; London agents: Macmillan and Co., Ltd., 1941.) 12*s*. 6*d*.

Professor Gamow has followed up his very successful *Birth and Death of the Sun* with a volume devoted to our own planet.

Starting from the evidence of radioactivity, the date of solidification of the oldest rocks is shown to be not more than two thousand million years ago (or two billion in the American terminology, which the author uses). The fact that it separated from the sun is next demonstrated, and the generally accepted hypothesis as to the nature of the close approach of two stars which led to this event is described.

Next comes the problem of how the moon was separated from the earth, and after that the question of how it got so far away. Then the arguments are marshalled for assuming that the Pacific Ocean represents the scar left when the moon was removed; the fact that the shape looks right is obvious to all of us, but the strongest single item of evidence lies in the fact that whereas everywhere else on the earth the basalt is covered by a granite layer 50 to 100 km. thick, yet in the Pacific Ocean no granite is found—the moon naturally took the whole of the top layer away.

More detailed consideration of the planets, their atmospheres, and the possibility of life on them brings us back to the detailed study of the earth's structure and interior. Here the hypothesis of continental drift is discussed, and isostasy is explained in a way which should be intelligible to the layman—no mean feat. This part of the book deals also of course with the processes of mountain building, ice ages and vulcanism. In connection with ice ages, the author rejects rather summarily the theory, which is widely held by many authorities, that variations in the amount of carbon dioxide or water in the atmosphere, by changing the absorption coefficient for radiation, may have been responsible for major changes of climate. He attributes the whole of the effects to changes in the earth's rotation round the sun, without even mentioning any of the arguments against this theory.

Chapter IX deals with life on the earth, and in particular with its origin. Plausible arguments are advanced for the theory that life might arise spontaneously, and it would be interesting to have the opinions of a biologist on the matter.

The book closes, unlike most biographies, with a glimpse into the future. As far as can be foreseen at present, there will " soon " be a period of much more intense mountain building. Then, about 50,000 A.D., there will be an ice age, when most of Europe and North America will be under thick ice, which will cover Oslo, Copenhagen, Stockholm and Leningrad, but not London, Paris or Berlin. Before this, however, the climate of the earth as a whole will be warmer, with its climax about 20,000 A.D. Many millions of years later again, the mountains will have been washed flat. It is suggested that the mammals, all increased in size, will still dominate the animal kingdom.

While this is happening, the moon may be expected to recede from the earth, until in 20 or 30 thousand million years she reaches her maximum distance, after which she will gradually approach the earth again. At the culminating point, the day and the month will be equal, each occupying 47 of our present days. All this, however, depends on the sun remaining much as at present. Professor Gamow, however, suggests that the sun may become a nova or a supernova in only ten thousand million years, in which case the earth and moon will be melted, and the subsequent history of the solar system will be simply a record of the gradual cooling of lifeless sun and planets. J. H. A.

The Human Meaning of Science, by ARTHUR H. COMPTON. Pp. xii + 88. (Chapel Hill: University of North Carolina Press, 1942.) 6*s*.

In the second of the three lectures in this volume Prof. Compton discusses the abandonment of determinism in modern physics and its bearing on human free will. " As one whose experiments have been partly responsible for this dramatic reversal of the physicist's point of view, I have been specially interested in tracing what the significance of this change may be to human life and thought." Those who have not yet realized the magnitude of the change of outlook will do well to ponder over this simple and outspoken lecture. Human freedom was in direct contradiction to the older system of physical laws ; either it must be an illusion or we must admit that somewhere in the organism a violation of the laws occurred. It is not in direct contradiction to the new quantum mechanics, though, as Compton points out, " this form of mechanics does not necessarily imply freedom ; by itself it implies chance ". Free will is no longer the antithesis of determinism ; it is the antithesis of chance. In its new form the problem of free will presents questions which require careful investigation. The lecture is mainly occupied with the justification of the new outlook, and little space is left to deal with the question, How can volition overcome chance ? I think the difficulty appears more formidable than it really is, because probability in statistical physics is commonly used in a loose way which would not be tolerated by statisticians. More particularly, in the general framework of equations of quantum theory (where there is no reference to special information), the probability is necessarily relative to assumed information which is not explicitly stated. From this point of view the setting aside of chance by volition is no more than the setting aside of *assumed*, but in this case erroneous, information that no volition is involved. Compton has some illuminating remarks on the relation of habit and free will, which may perhaps be summarized by describing habit as statistical volition.

The other lectures are on " Science, Religion and Human Growth " and " The Need for God in an Age of Science ". In the former the main thesis is that, whereas in historic times the human individual has changed little, and it is in the social characteristics of humanity that the chief growth has occurred, science is foremost in accelerating the social change. The latter is an appropriate reflection on the significance of the facts brought out in the two previous lectures. A. S. EDDINGTON

THE PROCEEDINGS OF
THE PHYSICAL SOCIETY

VOL. 54, PART 6 1 November 1942 No. 306

NEW DELHI
22. FEB. 1943
LIBRARY

INFLUENCE OF BARRIER DIMENSIONS ON THE VIBRATIONAL SPECTRUM OF ETHANE

By A. CHARLESBY, Ph.D., D.I.C.,
Imperial College of Science and Technology, S.W.7

MS. received 30 *March* 1942 ; *in revised form* 8 *May* 1942

ABSTRACT. Recent investigations on the internal rotation in ethane indicate the existence of a high potential barrier. A sinusoidal shape of barrier is assumed, and to obtain agreement with thermal measurements a height of 3000 cal./mole is required. In this paper the effect of a potential barrier of rectangular shape on the energy levels is investigated. In this case a barrier of height $V=1700$ cal./mole and width 1·35 radians leads to substantially similar energy levels to those predicted by the sinusoidal barrier previously considered. The effect of splitting in the energy levels is found to be slight for levels below V, and considerable for those above V. The energy levels for heavy ethane are also deduced.

It is concluded that the value previously deduced for the barrier height V, i.e. 3000 cal./mole, is only significant for a cosine form of variation. The evidence for this shape is by no means decisive.

§ 1. INTRODUCTION

RECENT investigations provide evidence for the existence of a potential which restricts rotation about the single C—C bond. Ethane has been chosen as a typical example, and the evidence arises mainly from thermal and spectroscopic results. Calculations based on the vibrational frequencies of the molecule lead to a variation in the specific heat with temperature, which may be compared with direct measurement.

The ethane molecule has eighteen modes of vibration, of which six are non-degenerate and twelve doubly-degenerate. Among recent assignments, that due to Crawford, Avery and Linnett (1938), and based on both infra-red and Raman spectra, is perhaps the most reliable. Of these vibrations, the twisting frequency is spectroscopically inactive, while the degenerate E'' or ${}^2\nu_\alpha{}^M$ frequency is ambiguous. Quite apart from zero-energy considerations, a full calculation is therefore impossible. An alternative procedure is therefore adopted, the uncertain frequency E'' being chosen to obtain agreement with the specific heat at room temperature. If free rotation about the C—C bond is assumed, a frequency of 750 cm.$^{-1}$ is required, while with highly restricted rotation a value of 1100 cm.$^{-1}$ is found. To distinguish between these two possibilities, the specific heat at low temperatures was investigated (Kistiakowsky and collaborators, 1935 to 1939: Hunsmann, 1938). The results obtained favoured the assumption of a high restricting barrier.

An alternative method of approach makes use of the heat of hydrogenation of ethylene, of which the most accurate determination is that of Kistiakowsky, Romeyn, Ruhoff, Smith and Vaughan (1935). This value may be used in conjunction with the known modes of vibration to calculate the equilibrium constant K_p in the reaction

$$C_2H_4 + H_2 \rightleftharpoons C_2H_6.$$

With this estimate, Teller and Topley (1935), and also Smith and Vaughan (1935), showed that no agreement could be obtained on the assumption of free rotation. Kemp and Pitzer (1936, 1937), however, showed that, with a high restricting barrier, agreement could be established, the value chosen (3000 cal./mole approximately) also accounting for the observed specific heat and entropy. Guggenheim (1941) verified that the agreement thus obtained between calculated and observed values of K_p was maintained over a wide range of temperature.

The cause of the restricting potential is by no means clear. Eyring (1932) estimated the interaction between electron pairs, and found a potential height of about 300 cal./mole. Penney (1934), using the method of molecular orbitals, concluded that at room temperatures, rotation about the C—C bond is by no means free.

A quantum mechanical investigation of the double rotator with one degree of rotational freedom was made by Nielsen (1932), and subsequently by Teller and Weigert (1933). The potential barrier assumed was a sinusoidal function of the angle between the methyl groups. The equation to be solved is of the form of Mathieu's equation:

$$\frac{d^2M}{dx^2} + (a + 16q\cos 2x)M = 0.$$

Kemp and Pitzer showed that, with a barrier of height 3150 cal./mole, agreement is obtained both with regard to specific-heat measurements and to the equilibrium constant. The shape of barrier assumed was criticized by Kistiakowsky and Wilson (1938) as being of arbitrary character. Pitzer and Kemp (1938) answered this objection, but failed to justify the form of barrier chosen.

Kistiakowsky, Lacher and Stitt (1939), from a study of the specific heat of ethane and heavy ethane, concluded that the existence of energy levels 0, 275 ± 10 cm.$^{-1}$, 520 ± 10 cm.$^{-1}$, and (very approximately) 725 cm.$^{-1}$ would account satisfactorily for the observed specific heat. These levels are equal to those deduced from a cosine form of barrier, of height 2750 cal./mole approximately. It was therefore concluded that these measurements afforded strong evidence for the shape of potential chosen, or at least for its degree of anharmonicity.

In this paper an attempt is made to discover whether different shapes of potential barrier cannot account equally well for the observed levels, which summarize the experimental conclusions. A rectangular type of barrier is chosen, and both height and width are varied. Satisfactory agreement can be obtained with a much lower level, and the evidence for a sinusoidal shape, or even for a corresponding anharmonicity, is therefore by no means decisive.

§ 2. DETERMINATION OF THE ENERGY LEVELS

The energy levels of the ethane molecule may be determined from the equation

$$\frac{\partial^2\Psi}{\partial\phi_1^2} + \frac{\partial^2\Psi}{\partial\phi_2^2} + \frac{8\pi^2 I}{h^2}[W_T - V]\Psi = 0. \qquad \ldots\ldots(1)$$

Here I is the moment of inertia of each methyl group, and ϕ_1, ϕ_2 are the corresponding angular co-ordinates about the common axis. V is the interaction energy between the end groups, and depends only on the angle $\phi_1 - \phi_2$ between them. W_T is the energy of the molecule, and the other symbols have their customary meanings.

To separate the solid-body rotation and torsional motion of the molecule, substitute

$$\xi = \frac{\phi_1 + \phi_2}{2}, \quad \eta = \phi_1 - \phi_2. \qquad \ldots\ldots(2)$$

Then (1) becomes

$$\frac{1}{2}\frac{\partial^2 \Psi}{\partial \xi^2} + 2\frac{\partial^2 \Psi}{\partial \eta^2} + \frac{8\pi^2 I}{\boldsymbol{h}^2}\{W_T - V(\eta)\}\Psi = 0.$$

The variables ξ, η are separable.

Let $\Psi = \psi(\xi)\chi(\eta)$, so that

$$\frac{1}{2}\chi\frac{\partial^2 \psi}{\partial \xi^2} + 2\psi\frac{\partial^2 \chi}{\partial \eta^2} + \frac{8\pi^2 I}{\boldsymbol{h}^2}\{W_T - V\}\chi\psi = 0,$$

$$\frac{1}{2\psi}\frac{\partial^2 \psi}{\partial \xi^2} = -\frac{2}{\chi}\frac{\partial^2 \chi}{\partial \eta^2} - \frac{8\pi^2 I}{\boldsymbol{h}^2}\{W_T - V\} = c, \text{ say,}$$

where c does not depend on χ, η.

The equation for ψ has a general solution:

$$\psi = Ae^{i\Lambda\xi} + Be^{-i\Lambda\xi}, \qquad \ldots\ldots(3)$$

where $\Lambda^2 = -2c > 0$.

This equation defines the rotational spectrum of the molecule, rotating bodily; this motion is quantized by the integer Λ.

The corresponding equation to determine the twisting energy levels can be written

$$\frac{\partial^2 \chi}{\partial \eta^2} + \frac{4\pi^2 I}{\boldsymbol{h}^2}\left(W_T - V - \frac{\Lambda^2 \boldsymbol{h}^2}{16\pi^2 I}\right)\chi = 0, \qquad \ldots\ldots(4)$$

or

$$\frac{\partial^2 \chi}{\partial \eta^2} + \frac{4\pi^2 I}{\boldsymbol{h}^2}(W - V)\chi = 0. \qquad \ldots\ldots(4\,a)$$

The quantity $W\left(= W_T - \frac{\Lambda^2 \boldsymbol{h}^2}{16\pi^2 I}\right)$ represents the energy due to the torsional motion alone, while $\frac{\Lambda^2 \boldsymbol{h}^2}{16\pi^2 I}$, the remaining part of the energy, is due to solid-body rotation.

Before proceeding to the solution of the equation for the internal twisting energy W, a boundary condition may be stressed. The wave function Ψ must not change if one methyl group is rotated by 2π, since the molecule does not thereby change, i.e.:

$$\Psi(\phi_1, \phi_2 + 2\pi) = \Psi(\phi_1, \phi_2).$$

However, if ϕ_2 is increased by 2π, η decreases by 2π, but ξ is only increased by π, so that

$$\psi(\xi)\chi(\eta) = \Psi(\phi_1, \phi_2) = \Psi(\phi_1, \phi_2 + 2\pi) = \psi(\xi + \pi)\chi(\eta - 2\pi).$$

But

$$\psi(\xi + \pi) = \psi(\xi)e^{i\Lambda\pi}.$$

Hence $$\chi(\eta-2\pi)=\chi(\eta)e^{-i\Lambda\pi}=\pm\chi(\eta), \qquad \ldots\ldots(5)$$

depending on whether Λ is even or odd.

The solutions of the equation

$$\frac{\partial^2\chi}{\partial\eta^2}+K^2(W-V)\chi=0, \quad \text{where } K^2=\frac{4\pi^2 I}{h^2}, \qquad \ldots\ldots(6)$$

therefore fall into two classes:

$$\chi(\eta) = \chi(\eta+2\pi); \qquad (\Lambda \text{ even}) \qquad \ldots\ldots(5\,a)$$

$$\chi(\eta) = -\chi(\eta+2\pi). \qquad (\Lambda \text{ odd}) \qquad \ldots\ldots(5\,b)$$

For simplicity, we put $\beta^2=K^2W$ and $\gamma^2=K^2(V-W)$, ...(6*a*) and (6*b*) assuming that $V>W$. The case where $V<W$ will be discussed later.

The potential energy V of interaction between the methyl groups is a function of the angle η between them. The shape of the barrier preventing free rotation about the C—C axis is unknown, and even its cause is not clear. It is, therefore, necessary to choose V as a function of η consistent with the known symmetry properties. Unlike preceding analyses, we assume a barrier of rectangular shape. In view of the three-fold symmetry of the rotators, V must repeat at intervals of $2\pi/3$ in η. Thus, if the barriers are each of width b, the valley between them must be of width a, where

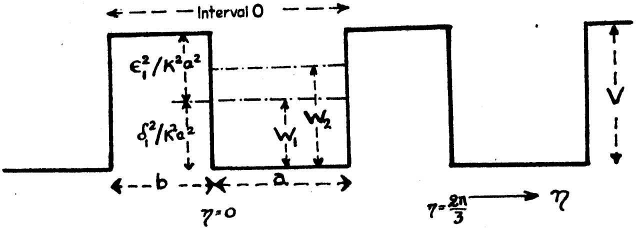


Figure 1. Variation of V with η.

$$a+b=2\pi/3. \qquad \ldots\ldots(7)$$

For convenience, each complete range of $2\pi/3$ is termed an *interval*, the interval of zero order extending from $-b$ to $+a$.

The function $\chi(\eta)$ must be so chosen that it satisfies equation (6) and fulfils the boundary condition $\chi(\eta+2\pi)=\pm\chi(\eta)$; moreover, both χ and $d\chi/d\eta$ must be continuous, even for discontinuous values of V. The form of equation (6) shows that if χ is a solution in a certain angular region η_0, and η_0 is increased by an integral number of times $2\pi/3$, χ and, more generally, $M\chi$ will be a solution in the new region, M being any quantity which is constant within the interval.

One can conveniently express χ for an angle η in the pth interval in terms of the value of χ for the corresponding angle η_0 in the zero interval, multiplied by an exponential chosen as so to fulfil the boundary condition (5)

$$\chi(\eta) = \chi(\eta_0)e^{i\alpha p}, \qquad \ldots\ldots(8)$$

$$\eta = \eta_0+p\,.\,2\pi/3 \qquad \ldots\ldots(8\,a)$$

and p is an integer. α is defined from the boundary condition

$$\chi(\eta_0+2\pi) = \pm\chi(\eta_0),$$

depending on whether Λ is even or odd. The simplest solutions are where

$$\begin{aligned} \alpha &= 0, \quad 2\pi/3, \quad 4\pi/3; \qquad (\Lambda \text{ even}) \\ \alpha &= \pi/3, \quad \pi, \quad 5\pi/3; \qquad (\Lambda \text{ odd}) \end{aligned} \qquad \ldots\ldots(9)$$

these solutions include all others.

The wave functions for the first three intervals can thus be expressed in terms of one of them, as shown in table 1. Only solutions for the interval 0 need therefore be found.

Table 1

	Λ even			Λ odd		
Interval (p)	$a=0$	$a=2\pi/3$	$a=4\pi/3$	$a=\pi/3$	$a=\pi$	$a=5\pi/3$
0 ($\eta=\eta_0$)	$\chi(\eta_0)$	$\chi(\eta_0)$	$\chi(\eta_0)$	$\chi(\eta_0)$	$\chi(\eta_0)$	$\chi(\eta_0)$
1 ($\eta=\eta_0+2\pi/3$)	$\chi(\eta_0)$	$\chi(\eta_0)e^{2\pi i/3}$	$\chi(\eta_0)e^{4\pi i/3}$	$\chi(\eta_0)e^{\pi i/3}$	$\chi(\eta_0)e^{i\pi}$	$\chi(\eta_0)e^{5\pi i/3}$
2 ($\eta=\eta_0+4\pi/3$)	$\chi(\eta_0)$	$\chi(\eta_0)e^{4\pi i/3}$	$\chi(\eta_0)e^{8\pi i/3}$	$\chi(\eta_0)e^{2\pi i/3}$	$\chi(\eta_0)e^{2\pi i}$	$\chi(\eta_0)e^{10\pi i/3}$

In the range $0<\eta_0<a$, $V=0$, so that (6) becomes

$$d^2\chi/d\eta_0{}^2+\beta^2\chi = 0,$$

with the solution

$$\chi(\eta_0) = Ce^{i\beta\eta_0}+De^{-i\beta\eta_0}.$$

The general solution, for an angle η in any interval p, is then found:

$$\chi(\eta)=\chi(\eta_0)e^{i\alpha p}=\{Ce^{i\beta\eta_0}+De^{-i\beta\eta_0}\}e^{i\alpha p}. \qquad \ldots\ldots(10)$$

Similarly, in the interval $-b<\eta_0<0$,

$$d^2\chi/d\eta^2-\gamma^2\chi = 0$$

has a solution,

$$\chi(\eta_0) = Ae^{\gamma\eta_0}+Be^{-\gamma\eta_0},$$

so that

$$\chi(\eta) = \chi(\eta_0)e^{i\alpha p}= (Ae^{\gamma\eta_0}+Be^{-\gamma\eta_0})e^{i\alpha p}. \qquad \ldots\ldots(11)$$

Continuity conditions

For continuity of χ on both sides of the boundary, $\eta = \alpha p$ (i.e. $\eta_0= 0$),

$$C+D=A+B, \qquad \ldots\ldots(12\,a)$$

and continuity in $d\chi/d\eta$ at the same boundary implies the condition

$$A\gamma-B\gamma = Ci\beta-Di\beta. \qquad \ldots\ldots(12\,b)$$

Continuity of χ at the boundary when $\eta_0=a$ ($\eta= -b+p.\,2\pi/3 = a+(p-1)2\pi/3$) implies

$$\chi(a)e^{i\alpha(p-1)}= \chi(-b)e^{i\alpha p},$$

or

$$Ce^{i\beta a}+De^{i\beta a}= (Ae^{-\gamma b}+Be^{\gamma b})e^{i\alpha}, \qquad \ldots\ldots(12\,c)$$

while continuity in $d\chi/d\eta$ at the same boundary leads to the relation

$$Ci\beta e^{i\beta a}-Di\beta e^{-i\beta a}= e^{i\alpha}(A\gamma e^{-\gamma b}-B\gamma e^{\gamma b}). \qquad \ldots\ldots(12\,d)$$

These four equations will only have roots ($\neq 0$) if the following relation holds between the four parameters A, B, C and D:

$$\frac{\gamma^2-\beta^2}{2\beta\gamma}\sinh\gamma b\sin\beta a+\cosh\gamma b\cos\beta a=\cos\alpha. \qquad \ldots\ldots(13)$$

The permitted energy levels may be determined from the solutions of this equation.

To solve the equation for β and γ, the following substitutions may be made:

$$\left.\begin{array}{llr} \beta a=\delta, & \gamma a=\epsilon, & y=\gamma/\beta=\epsilon/\delta, \\ K^2Va^2=H^2, & \multicolumn{2}{l}{s=\cos\alpha\,(=1,\ \tfrac{1}{2},\ -\tfrac{1}{2},\ -1).} \end{array}\right\} \qquad \ldots\ldots(14)$$

The quantity H^2 determines to a considerable extent the effect of the barrier, while δ^2 and ϵ^2 are respectively proportional to the distance from an energy level to the base and to the top of the potential barrier. The following relations hold:

$$\left.\begin{array}{c} \delta^2+\epsilon^2=H^2, \\ \delta^2=H^2/(1+y^2), \\ \delta=\epsilon y, \end{array}\right\} \qquad \ldots\ldots(15)$$

while the fundamental equation (13) becomes

$$\frac{y^2-1}{2y}\sinh\gamma b\sin\delta+\cosh\gamma b\cos\delta=s. \qquad \ldots\ldots(16)$$

The value of K^2, assuming a tetrahedral angle at C, a C—H spacing of 1·09 A. and a H mass of $1{\cdot}673\times 10^{-24}$ gm., is found to be $4{\cdot}78_8\times 10^{14}$.

Pitzer, assuming a cosine form of barrier, obtains a height of about 3000 cal./mole ($2{\cdot}14\times 10^{-13}$ ergs or 0·135 ev.). We shall at first deal with barriers above 2000 cal./mole, and deduce the corresponding frequencies. To simplify the calculations we consider primarily those energy levels which lie well below the potential barrier. For these levels we make take $\epsilon \overset{\sim}{>} \delta$. Pitzer's value corresponds to a value of $103\,a^2$ for $K^2Va^2(=\epsilon^2+\delta^2)$, so that $\epsilon \geqslant 7a$. Physical conditions would lead us to assume that a and b are of the same order. Then $\gamma b\sim\gamma a=\epsilon$, so that a probable minimum value for $\sinh\gamma b$ and $\cosh\gamma b$ is about 500. Compared with this figure, the splitting factor s is small, and may be neglected in a first approximation. Detailed calculation will show how far this approximation is valid.

At the same time, we may take $\sinh\gamma b$ and $\cosh\gamma b$ as equal. The equation can now be written

$$\left(\frac{y^2-1}{2y}+\cot\delta\right)\sin\delta\sinh\gamma b=0. \qquad \ldots\ldots(17)$$

Approximate solutions occur for values of δ which satisfy the equation

$$\frac{1-y^2}{2y}=\cot\delta \qquad \ldots\ldots(18)$$

when $$\delta^2=H^2/(1+y^2).$$

A simple graphical method of solution which shows the relation between the roots can be obtained by calculating δ for various values of y (a suitable value of H^2 being assumed). Graphs of $(1-y^2)/2y$ and $\cot\delta$ are then plotted against δ, points of intersection giving the required roots. Different sets of solutions can be obtained by choosing different values for H^2. For any given value of y, the only effect of this change is to multiply δ by a constant factor, and hence to contract or expand the $(1-y^2)/2y$ curve along the δ axis. The $\cot\delta/\delta$ curve of course remains unchanged. It is therefore possible to draw up a series of solutions δ_1, δ_2, δ_3 ... determined in terms of the parameter H.

Several interesting conclusions emerge from these results. In the first place δ depends only on H, except, perhaps, for low values of the latter, when the approximations considered may break down. Thus, if H remains constant, the ratio $W/V = \delta^2/H$ will not change. Increases in the barrier height, accompanied by suitable changes in the barrier width, will then cause proportionate increases in the energy levels, the relative heights of the permitted energy levels remaining unaltered.

In the second place, variations in the barrier constant H over a wide range produce only small changes in δ. It is therefore not very satisfactory to deduce H from data on δ.

Finally, the relative values of δ_1, δ_2, δ_3 are very nearly in the ratios 1 : 2 : 3, ... (figure 2), and depart from these ratios by only a few per cent when H^2 reaches values as low as 50. Since it is these relative values which determine the relative heights of the energy levels W_1, W_2, W_3, ..., one cannot alter the latter greatly by changing the value of the barrier constant H, determined in terms of V and a.

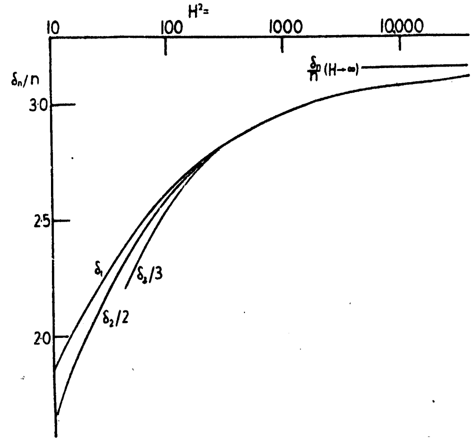


Figure 2. δ_n/n as a function of H^2.

A possible exception may occur when the energy levels considered approach the barrier height. In this case the equation from which they are deduced would in any case require re-examination.

Alternative method

In the case considered above, when $\gamma b \gg 1$, it is possible to solve equation (17) very accurately by the following substitution.

Let $$y = \gamma/\beta = \epsilon/\delta = \tan\theta. \qquad \text{......(19)}$$

Then $$\frac{y^2-1}{2y} = -\cot 2\theta, \qquad \text{......(19 a)}$$

and $$\cot\delta-\cot 2\theta = s/(\sin\delta\sinh\gamma b)\,; \qquad \ldots\ldots(20)$$

moreover $$\delta^2=H^2/(1+\gamma^2)=H^2\cos^2\theta, \quad \text{i.e. } \delta=\pm H\cos\theta. \qquad \ldots\ldots(20\,a)$$

(This assumes $\sin\delta\neq 0$ and $\sinh\gamma b \approx \cosh\gamma b$.)

When the splitting factor s is neglected, solutions of (20) are given by $\delta=2\theta-n\pi$. It is most convenient to choose any suitable value of δ, and calculate the value of H such that $H=\pm\delta\sec\theta$, where $\theta=(\delta+n\pi)/2$. Barrier constants H corresponding to a series of known energy levels (as determined by δ) can then be readily deduced.

An ambiguity arises, since two values of H can be obtained:

$$H=\pm\delta\sec\delta/2 \quad \text{and} \quad H=\pm\delta\sec(\delta/2+\pi/2). \qquad \ldots\ldots(21)$$

Only one of these is correct; the other arises from the two possible values $\pm\gamma$, $\pm\beta$, only the squares of which are determined by the given data. Thus if γ and β are of the same sign, only values of θ in the first and third quadrants can be used, while if one chooses γ, β of opposite sign, θ is in the second or fourth quadrant. These two possibilities are included in the $\cot 2\theta$ values, and it becomes necessary to separate the correct solution when inserting in the value for H. If γ, β are both chosen positive, only those values of θ need be considered which lie in an odd quadrant ($\ldots -1, 1, 3, \ldots$).

It is also convenient to choose the sign in (20 a) so that H, δ are always positive. This entails the + sign when θ is in quadrants $-3, 1, 5, \ldots$ and the − sign when θ is in quadrants $-1, 3, 7, \ldots$.

§ 3. CALCULATION OF ENERGY LEVELS

For any given value of the barrier constant H, it is possible to determine the relative values of the energy levels, measured from the ground level W_1. To each solution δ_n (or β_n) of (13) or (17) corresponds a permitted level,

$$W_n=\beta_n^2/K^2=\delta_n^2/K^2a^2=\delta_n^2V/H^2, \qquad \ldots\ldots(22)$$

so that $$\frac{W_n-W_1}{W_2-W_1}=\frac{\delta_n^2-\delta_1^2}{\delta_2^2-\delta_1^2}\,; \qquad \ldots\ldots(23)$$

all energy levels being measured relative to the ground level W_1.

The energy levels of the ethane molecule may be determined from measurements of the specific heat. Kistiakowsky, Lacher and Stitt (1938 and 1939) have shown that energy levels of 0, 275, 520, 725 … cm.$^{-1}$ account satisfactorily for the observed heat capacity. It is proposed to determine a in terms of H, so that the first energy level (W_2-W_1) shall be that (275 cm.$^{-1}$) observed. Higher levels will then be calculated and compared with those obtained experimentally. At the same time the corresponding value of V may be deduced.

We require a, such that W_2-W_1 or $(\delta_2^2-\delta_1^2)/K^2a^2=275$ cm.$^{-1}=5{\cdot}45\times 10^{-14}$ ergs; hence $a=(\delta_2^2-\delta_1^2)/26{\cdot}1$. Since δ_1, δ_2 are known for any value of H within the range under consideration, a may thus be evaluated in terms of H, so that, whatever the value of H, the first energy-level is always 275 cm.$^{-1}$. The values of a and b/a $(=r)$ corresponding to different barrier constants H are shown in figure 3 a. Except for comparatively shallow barriers, the variation in a is slight. It is of interest to note that the value of a obtained is of the expected order ($\pi/3$). An increase of the first energy-level difference from 275 cm.$^{-1}$ to less than 400 cm.$^{-1}$ leads to unacceptable solutions ($a > 2\pi/3$). The corresponding potential height V is obtained from the relation

$$V = H^2/K^2a^2.$$

The higher energy levels are determined by means of (23), which may now be written

$$W_n - W_1 = \{(\delta_n{}^2 - \delta_1{}^2)/(\delta_2{}^2 - \delta_1{}^2)\} . 275 \text{ cm.}^{-1}. \qquad \ldots\ldots(24)$$

For large values of H ($H^2 > 100$), $\delta_n \propto n$, so that $W_3 - W_1 = 733$ cm.$^{-2}$ This is in direct contradiction with the observed value of about 520 cm.$^{-1}$ (± 10 cm.$^{-1}$). Similarly, the calculated third energy-level difference ($W_4 - W_1$) 1375 cm.$^{-1}$ is far greater than that (725 cm.$^{-1}$) obtained from experiments.

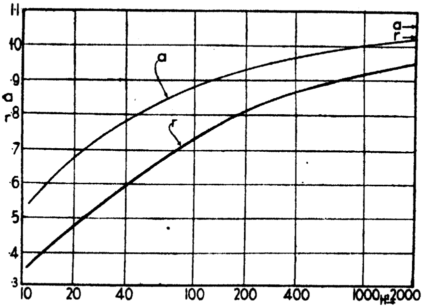


Figure 3 *a*. Relation between a, a/b ($=r$), and H^2 (first energy level = 275 cm.$^{-1}$).

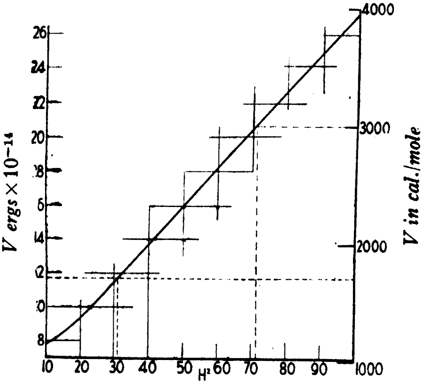


Figure 3 *b*. Relation between V and H (first energy level = 275 cm.$^{-1}$).

For smaller values of H, $W_3 - W_1$ decreases slowly until a value $H^2 \sim 40$ (2000 cal./mole) is reached, when the energy levels considered are of the same height as the potential barrier, and the equation used requires consideration. Even in this case, however, the second energy-level difference does not lie below 600 cm.$^{-1}$

We must therefore conclude that, with the shape of the barrier chosen, and for potential barriers which are not very much smaller than those obtained by Pitzer, no satisfactory agreement can be obtained between the calculated energy levels (with the exception of the first) and those found experimentally. Moreover, for barriers greater than 4000 cal./mole, and of width of about $\pi/3$, the lowest energy levels found are well-nigh independent of barrier height. The energy levels for smaller values of the barrier constant H will be considered below.

§ 4. EFFECT OF SPLITTING FACTOR (s)

In the above work, the influence of s has been neglected, and the solutions obtained are only approximate in nature. The effect of s on the energy levels will be evaluated in two cases, depending on how close to the top of the potential barrier the energy level considered lies. These two cases correspond to $\gamma \gg 1$ and $\gamma \ll 1$. The intermediate case is less amenable to direct treatment.

(*a*) $\gamma \gg 1$, so that $\cosh \gamma b = \sinh \gamma b$.

This implies that the difference $V-W$ between potential-barrier height and energy level considered is not very small.

The equation $\quad -\cot 2\theta \sin \delta \sinh \gamma b + \cos \delta \cosh \gamma b = s$

can, subject to the assumed condition, be written

$$\sin (2\theta - \delta) = s \sin 2\theta / \sinh \gamma b. \qquad \ldots\ldots(25)$$

The approximate solutions previously considered correspond to the values $\delta = 2\theta - n\pi$. To obtain more exact solutions, neighbouring values of θ are chosen, and the corresponding values of δ determined from the condition $\delta = \pm H \cos \theta$, which must be accurately obeyed. γb is then calculated from δ, H and the corresponding ratio b/a. Both sides of the equation are then obtained and compared. The solutions for several values of H are shown in table 2. The splitting of the third level (when $H^2 = 63{\cdot}1$) and of the fourth level (when $H^2 = 100$) shows that the effect is slight even when the energy levels are not far removed from the top of the potential barrier. For levels well below this, the splitting becomes negligible, due to the considerable value of $\sinh \gamma b$.

The apparent ambiguity in θ previously mentioned leads to an interesting alternation in the split levels. Only those values of θ which lie in an odd quadrant are acceptable (γ, β positive), so that $\sin 2\theta$ is always positive. Consider two consecutive levels, corresponding to values of δ differing by about π. Then if $2\theta = \delta_1$ is one acceptable solution (θ in odd quadrant), and δ_1 is increased by about π to δ_2, 2θ must be increased not by π, but 2π, if θ is still to remain in an odd quadrant. Thus in successive levels $\sin (2\theta - \delta)$ will reverse sign, whereas the sign of $s \sin 2\theta / \sinh \gamma b$ is unaffected. The effect of a given splitting factor s will, in alternate levels, be to increase or decrease the energy levels. This appears more clearly in the limiting case considered below.

Table 2. Split levels ($\delta < H$)

H^2	b (approx.)	$s=$ 1	$\frac{1}{2}$	0	$-\frac{1}{2}$	-1
100	2·5–2·8	$9{\cdot}84_2$	$9{\cdot}82_7$	$9{\cdot}81_2$	9·80	$9{\cdot}78_8$
63·1	5·2	$7{\cdot}16_7$	$7{\cdot}17_0$	$7{\cdot}17_2$	$7{\cdot}17_5$	$7{\cdot}17_8$

(*b*) $\gamma \ll 1$.

The energy levels considered are nearly equal to, but are still less than, the potential barrier height V. The case when they are greater than V will be dealt with later. The approximation $\sinh \gamma b = \cosh \gamma b$ is no longer valid. Equation (13) can be simplified ($\gamma^2 \ll \beta^2$ $(\gamma b)^2 \ll 1$) and becomes

$$-\frac{\delta}{2} \cdot \frac{b}{a} \sin \delta + \cos \delta = s. \qquad \ldots\ldots(26)$$

The values for δ_n, when this is nearly equal to H, are $(n-1)\pi$ approximately. Let $\delta_n = (n-1)\pi + f$, and assume $\sin f \sim f$, $\cos f = 1$. Then replacing in (26) and simplifying $\quad f = 2\{1 - s(-1)^{n-1}\} . a/b(n-1)\pi$,

$$\delta_n = (n-1)\pi + \frac{2a}{b(n-1)\pi} \{1 - s(-1)^{n-1}\}. \qquad \ldots\ldots(27)$$

The factor $\{1-s(-1)^{n-1}\}$ produces an alternation in the split levels, depending both on the level considered n and on the solid-body rotation quantum number Λ. Table 3 shows the maximum splitting for levels below the potential barrier height. The values obtained are, however, only approximate, since f is not sufficiently small to justify the approximation made. Moreover, the split values of δ_n thus obtained do not relate to a single value of H, but to that value of H for which the energy level equals the barrier height. The actual splitting is appreciably smaller.

Table 3. Maximum splitting ($\delta^2 \ll H^2$)

H (approx.)	n	a/b	$\dfrac{2a}{b(n-1)\pi}$	δ					
				Λ even			Λ odd		
3·2	2	0·31	0·2	3·54	3·24	3·24	3·14	3·44	3·44
6·4	3	0·59	0·188	6·28	6·56	6·56	6·65	6·37	6·37
9·6	4	0·71	0·15	9·72	9·50	9·50	9·42	9·64	9·64

§ 5. SPECIAL CASES

A number of assumptions have been made as to the relative importance of different factors. These may now be examined to discover whether the approximations made are valid under the conditions concerned.

In the general solution we have taken $\cosh \gamma b = \sinh \gamma b$. A more accurate analysis, but still neglecting the splitting factor, leads to the equation

$$\frac{1-y^2}{2y} = \cot\delta \coth\gamma b.$$

In the graphical solution the $\{1-y^2/2y, \delta\}$ curve will remain unaltered, while the $\{\cot\delta, \delta\}$ curve will have ordinates reduced by a variable factor $\coth\gamma b$. Since b is of the order of unity, the correction will only be appreciable for values of γ of about 2 or less. Only energy levels very nearly equal to the potential-barrier height will be affected. Moreover, if $\gamma b = 1$, the ordinates of the $\{\cot\delta, \coth\gamma b\}$ curve are decreased by 24 %. The solutions of the modified equation will then change, but by an amount considerably less than this, due to the steepness of the $\{\cot\delta, \delta\}$ curve. Thus, even in this extreme case, where $W/V > 0{\cdot}9$, the values of δ will differ from those obtained from the simpler equation $(1-y^2)2y = \cot\delta$ by less than 2 %. We are therefore justified in neglecting the $\coth\gamma b$ correction term in all but the most exceptional cases.

In certain cases the splitting factor s ($\neq 0$) may produce appreciable changes in the energy levels. These generally occur when the value of $\cosh\gamma b$ and $\sinh\gamma b$ are sufficiently small, i.e. for energy levels nearly equal to the potential height V. These have already been considered. A more special case occurs when in the general equation

$$\frac{y^2-1}{2y}\sin\delta\sinh\gamma b + \cos\delta\cosh\gamma b = s$$

(where γb is quite large) the left-hand side is small because both $\cos\delta$ and $(y^2-1)/2y$ are nearly zero. In this way the left-hand side may become comparable to unity in spite of the large values of $\sinh\gamma b$, $\cosh\gamma b$; s is no longer negligible, and may then exert an appreciable effect on energy levels which are by no means comparable with V.

It may be readily shown that the effect of the splitting factor s is to change δ by a very small amount, of the order of $s/\sinh\gamma b$. The splitting effect of s does not therefore differ appreciably from the general case.

§ 6. ENERGY LEVELS FOR LOW BARRIERS

Until the present moment we have limited the investigation to barriers of height greater than that defined by $H^2=40$, and to energy levels below the barrier height. This is sufficient to show that a 3000 cal./mole barrier can only succeed in correlating the experimental data if, in addition, a cosine shape of potential function is assumed. We shall now extend the work to lower barriers, and consider whether a suitable choice of potential height and breadth cannot be found to afford equally good agreement. Moreover, the splitting produced by the threefold degeneracy of the system will likewise be investigated.

For levels whose height W exceeds the potential-barrier height V, the difference $V-W$ or γ^2 is negative. Putting $\gamma^2=-\lambda^2$ in equation (13) and simplifying gives the alternative formula:

$$-\frac{\lambda^2a^2+\delta^2}{2\lambda a\,.\,\delta}\sin\lambda b\,.\,\sin\delta+\cos\lambda b\cos\delta=s. \qquad \dots\dots(28)$$

Solutions of this equation are not readily found as previously. A number of special solutions were obtained by putting $\delta=n\pi$, $(n+\frac{1}{2})\pi$, etc. The values of λa could then be obtained for any desired value H^2,

$$\delta^2-(\lambda a)^2=H^2, \qquad \dots\dots(29)$$

and the corresponding values λb deduced from the known ratio b/a. Since this ratio is obtained from the first two energy levels, whose values have only been obtained for $H^2>9{\cdot}8$, the calculations have been limited to greater values of H than 3·1.

Consider the variation of the function for $\delta=2\pi$. The equation then becomes $\cos\lambda b=s$. Solutions neglecting splitting ($s=0$) are $\lambda b=\pi/2$, $3\pi/2$, etc., and from these values the corresponding values H can be obtained. The split levels for $s=\pm\frac{1}{2}$ fall on either side of the non-split level, ($\lambda b=\pi/3$, $2\pi/3$, etc.), while for $s=\pm 1$ the split levels from adjacent levels overlap completely ($\lambda b=0$, π, 2π, etc.).

To discover whether this overlap occurs for all values of δ, the graph of the function

$$s=-\frac{(\lambda a)^2+\delta^2}{2\lambda a\,.\,\delta}\sin\lambda b\sin\delta+\cos\lambda b\cos\delta \qquad \dots\dots(30)$$

can be considered as a function of δ, keeping H constant. Two typical cases ($H^2=14$, $H^2=31$) are shown in figure 4. The values of δ can be directly obtained for values of s equal to 1, $\frac{1}{2}$, $-\frac{1}{2}$, -1. It will be seen that the degeneracy produced at $\delta=2\pi$ is only exceptional. The large values assumed by s in the transition from $\delta > H$ to $\delta < H$ are also clear. Nevertheless the splitting effect of s is very marked for all the range considered, and it becomes difficult to correlate the assembly of six split levels with a single level as required for the experimental results. We have taken the average value of a level as that corresponding to $s = 0$. For $H^2 = 31$, the mean energy levels relative to the ground level are shown below (table 4), as well as the split levels from which they arise.

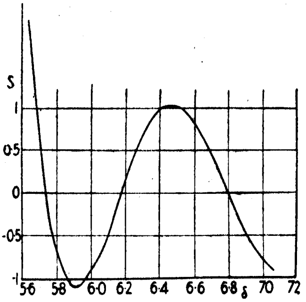


Figure 4 *a*. Plot of s as function of δ ($H^2 = 31$).

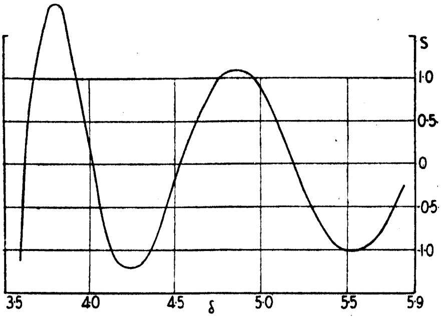


Figure 4 *b*. Plot of s as function of δ ($H^2 = 14$).

Table 4. Energy levels for 1700 cal./mole; 1·35 radians barrier ($H^2 = 31$)

Energy levels (cm.$^{-1}$) relative to ground level									
First	Second	Third				Fourth			
0	275	523				621			
Splitting negligible	negligible	511	516	531	549	679	645	602	576
Splitting factor		1	$\frac{1}{2}$	$-\frac{1}{2}$	-1	1	$\frac{1}{2}$	$-\frac{1}{2}$	-1

When Λ is even, each level is tripled ($s=1$, $s=\frac{1}{2}$, $s=\frac{1}{2}$); when Λ is odd, the levels correspond to $s=-1, \frac{1}{2}, \frac{1}{2}$.

The logarithmic plot of $\log\delta$ against $\log H$ for levels above V is almost linear. Over the region $12 < H^2 < 60$ the deviation from this law is of the order of several per cent. Then we may put approximately

$$\delta = CH^m,$$

where C, m are constants which may be determined for each level and have the approximate values

third level:	$C = 1{\cdot}22_7$	$m = 0{\cdot}89_8$
fourth level:	$C = 1{\cdot}65_2$	$m = 0{\cdot}76_8$
fifth level:	$C = 2{\cdot}11_3$	$m = 0{\cdot}68_0$

Over the range of values under consideration the first two levels W_1, W_2 remain less than V, and δ_1, δ_2 may hence be determined as before. This suffices to determine a as a function of H. Further solutions for δ_3, δ_4 may then be obtained as above (figure 5), and the corresponding energy levels $W_3 - W_1$, $W_4 - W_1$, etc., calculated. Best agreement with observed values is obtained for $H^2 = 31$. In table 5 the energy levels are calculated for various rectangular

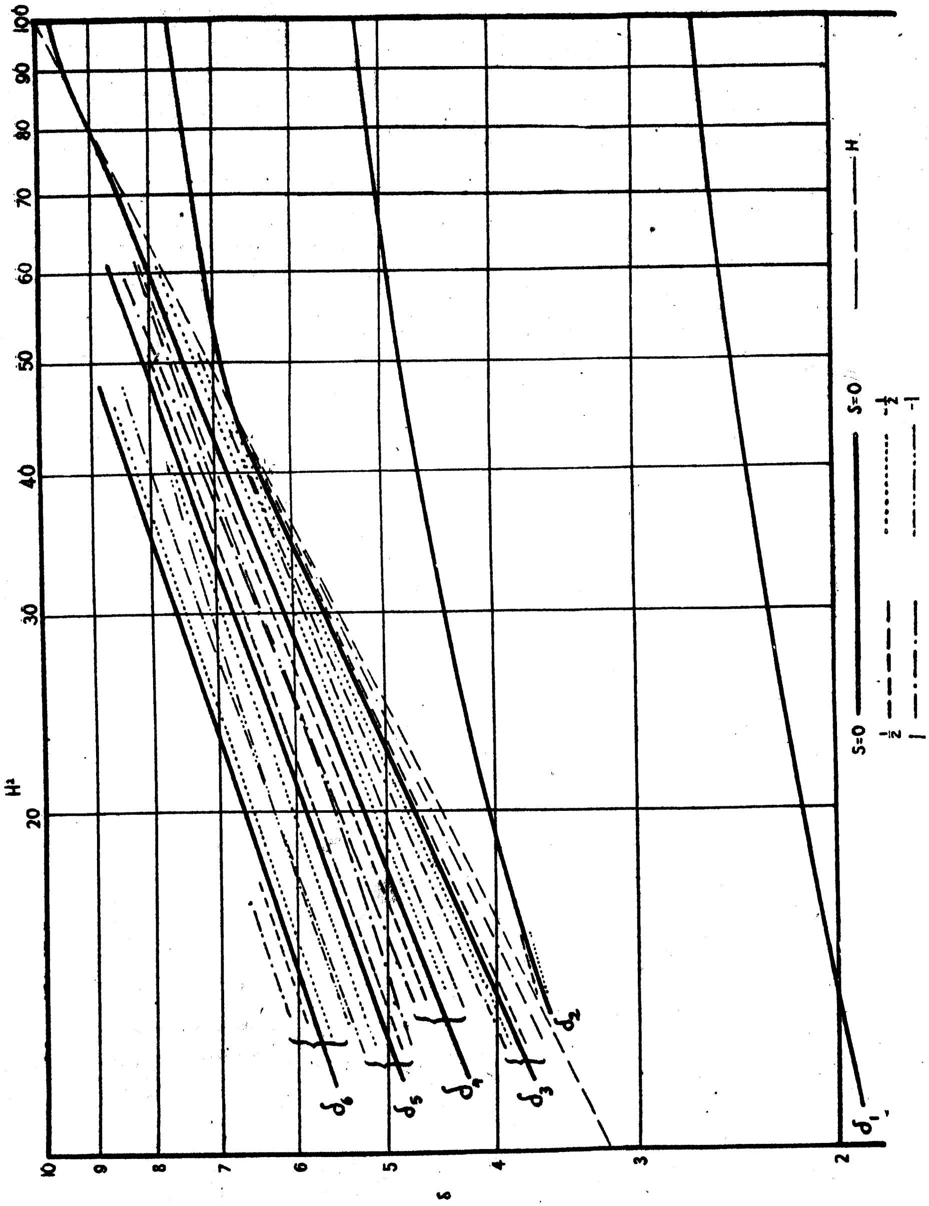


Figure 5. δ as a function of H (a adjusted so that $W_2 - W_1 = 275$ cm.$^{-1}$).
If $W \gtrsim V$, energy-level splitting shown. Λ even: $s = 1, -\frac{1}{2}, -\frac{1}{2}$.
Λ odd: $s = -1, \frac{1}{2}, \frac{1}{2}$.
If $W < V$, energy-level splitting is inappreciable.

barriers. In the case of the 3000 cal./mole barrier, the height is that deduced by Pitzer for a cosine barrier, while the width is chosen to reproduce the 275 cm.$^{-1}$ level. It will be noted that the other levels are then not reproduced satisfactorily.

Table 5. Comparison of energy-level differences (rectangular barrier)

H^2	Barrier height (cal.)	Barrier width (radians)	Energy level (cm.$^{-1}$)	Levels above ground (cm.$^{-1}$)	
∞	∞	1·03	275	733	1357
71·5	3000*	1·25	275	704	
31	1700	1·35	275	523	621
29	1630	1·35	275	503	596
†	—	—	275±10	520±10	725± ?

* Value obtained by Pitzer (cosine barrier).
† Value obtained by Kistiakowsky and colleagues (specific-heat measurements).

The effect of Λ on the torsional energy levels is shown in figure 6 for the case of $H^2=31$ (1700 cal./mole barrier). The energy levels W are, of course, proportional to δ^2. The mean values for any level (i.e. assuming $s=0$) are shown in broken lines.

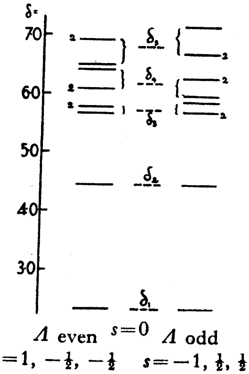


Figure 6. Splitting of levels for $H^2=31$. (δ_1, δ_2, splitting is inappreciable) δ_3, δ_4, δ_5 consist of one singlet and one doublet.

§ 7. APPLICATION TO C_2D_6

From the values of V and a deduced for ethane, it is possible to calculate the levels which should be observed in hexadeuteroethane, assuming the barrier constants to remain unaltered. The only change is then due to increased inertia, $I_D=2I_H$, so that $K_D{}^2=2K_H{}^2=9{\cdot}57_6\times10^{14}$ and $H_D{}^2=2H_H{}^2$.

It was found above that satisfactory agreement could be obtained for $H_H{}^2=31$ in the case of ethane. The energy levels for $H_D{}^2=62$ should therefore be investigated. The values of δ obtained as solutions of (13) are given in table 6. The corresponding energy levels W are then deduced from the expression

$$\delta^2=K_D{}^2Wa^2,$$

where K_D^2 has the value given above, and a is to be taken equal to the value deduced for the case of ethane ($H_H^2=31$, $a=0{\cdot}74_5$). As a check, the levels for ethane were deduced at the same time. The values thus obtained differ by about 1 cm.$^{-1}$ from those obtained more directly.

Table 6

$V=1700$ cal./mole $b=1{\cdot}35$ $a=0{\cdot}745$

C_2D_6 ; $H^2=62$			C_2H_6 ; $H^2=31$		
δ	Energy (cm.$^{-1}$)	Energy levels * (cm.$^{-1}$)	δ	Energy (cm.$^{-1}$)	Energy levels * (cm.$^{-1}$)
2·50	59	0	2·29	99	0
$4{\cdot}9_3$	231	172	4·43	373	274
$7{\cdot}1_2$	481	432	$5{\cdot}73_5$	625	524

* Values relative to ground level.

§ 8. CONCLUSION

The height of the potential barrier restricting free rotation in ethane has previously been evaluated assuming that the variation with angle of torsion follows a sinusoidal law, in which the only parameter is therefore the height V. If the maximum variation in V is taken as about 3000 cal./mole, satisfactory agreement is established between theoretical and experimental results. The lowest energy levels of the torsional vibration which are obtained from specific-heat measurements agree well with those deduced theoretically from the 3000 cal./mole barrier.

If, however, a barrier of the same height but of different (rectangular) shape be assumed, the energy levels no longer agree satisfactorily with experimental evidence. Nor can this agreement be obtained by variation in the barrier dimensions (height and width) except for the considerably lower value 1700 cal./mole (barrier width = 1·35 radians). Agreement is then again obtained with the experimental values, although the fourth energy level is somewhat less satisfactory than for the cosine functions. This could probably be improved, however, by small variations in the initial energy levels used in the calculations, since these are only known approximately. Moreover, the effect of splitting of the higher levels is found to be very considerable.

The energy levels for the heavy-ethane molecule have been calculated on the basis of the barrier dimensions found suitable in the case of ethane. Comparison with the values deduced by Kistiakowsky and colleagues is, however, impossible, since these have not been published. Barriers below 1100 cal./mole in the case of ethane have not been investigated.

It appears from these results that there is no decisive evidence for a cosine form of barrier. Moreover, the estimated height (3000 cal./mole) is only significant in so far as a barrier of this shape is assumed. The figure obtained should therefore only be treated as a convenient means of comparing data for different molecules.

More detailed experimental evidence is therefore required before an accurate estimate of the shape and dimensions of the potential barrier can be obtained.

§ 9. ACKNOWLEDGEMENT

The writer wishes to thank Assistant Professor W. G. Penney, D.Sc., for advice and criticism.

REFERENCES

Crawford, B. L., Avery, W. H. and Linnett, J. W., 1938. *J. Chem. Phys.* **6,** 682.
Eyring, H., 1932. *J. Amer. Chem. Soc.* **54,** 3191.
Guggenheim, E. A., 1941. *Trans. Faraday Soc.* **37,** 97, 271.
Hunsmann, W., 1938. *Z. phys. Chem.* B, **39,** 1.
Kemp, J. D. and Pitzer, K. S., 1936. *J. Chem. Phys.* **4,** 749 ; 1937. *J. Amer. Chem. Soc.* **59,** 276 ; 1938. *Ibid.* **60,** 1515.
Kistiakowsky, G. B., Lacher, J. R. and Stitt, F., 1938. *J. Chem. Phys.* **6,** 407 ; 1939. *Ibid.* **7,** 289.
Kistiakowsky, G. B. and Nazmi, F., 1938. *J. Chem. Phys.* **6,** 18.
Kistiakowsky, G. B., Romeyn, J., Ruhoff, J. R., Smith, H. A. and Vaughan, W. E., 1935. *J. Amer. Chem. Soc.* **57,** 876.
Kistiakowsky, G. B. and Wilson, E. B., 1938. *J. Amer. Chem. Soc.* **60,** 495.
Nielsen, H. H., 1932. *Phys. Rev.* **40,** 445.
Penney, W. G., 1934. *Proc. Roy. Soc.* A, **144,** 166.
Smith, H. A. and Vaughan, W. E., 1935. *J. Chem. Phys.* **3,** 341.
Teller, E. and Topley, B., 1935. *J. Chem. Soc.* 876.
Teller, E. and Weigert, K., 1933. *Nachr. Ges. Wiss. Göttingen,* p. 218.

THE THERMAL EXPANSION OF ALUMINIUM: FURTHER EXPERIMENTS

By A. J. C. WILSON, Ph.D.,
Cavendish Laboratory, Cambridge

MS. received 20 *May* 1942

ABSTRACT. It is found that the double curvature previously reported in the thermal expansion coefficient of aluminium is characteristic of unannealed filings. After the filings have been heated above 600° c., it disappears. The following values are obtained for the expansion coefficient of annealed filings :

t	0°	100°	200°	300°	400°	500°	600° c.
$10^6\alpha$	22·8	24·0	26·0	28·3	30·3	32·7	35·0

Certain observations on the effect of cold-work are recorded.

§ 1. INTRODUCTION

The writer's previous measurements of the expansion of aluminium (Wilson, 1941) were somewhat severely criticized in a review in the *Metallurgist* (27 June 1941). The chief grounds for the criticism were (*a*) that the results were not in accord with those of previous work, and (*b*) that no details of the annealing treatment were given. The first objection is not very important, particularly as the two determinations specifically cited as in good agreement (Taylor, Willey, Smith and Edwards, 1938; Esser and Eusterbrock, 1940–41) differ by about 7 % at 500° c.; it was, however, thought desirable to investigate the effect of annealing.

It has usually been considered that cold-working has little effect on the properties of pure aluminium. Tammann (1934), after describing changes produced by cold-working in the electrical resistance, thermal e.m.f. and other properties of silver, gold and copper, states that aluminium behaves differently in that the electrical resistance and thermal e.m.f. are unchanged, and that the x-ray reflections are not broadened. Wood (1939) found that moderate cold-working increased the lattice parameters of silver, copper, iron, nickel and molybdenum; extreme cold-working reduced them below those of the undistorted metals. Aluminium, on the other hand, was "spontaneously self-recovering during cold-working, and for this reason was found not to exhibit the changes described for copper and silver". Thomassen and Wilson (1933) found, however, that reflections from aluminium worked at the temperature of solid CO_2 were broader than reflections from aluminium worked at room temperature.

In order to see if the writer's S-shaped (α, T) curve could be explained by lack of annealing, the spacing measurements on which it was based were divided into two sets: those taken before and those taken after the aluminium had been heated above 600° c. It was found that the former set showed the double curvature; the latter did not, but there was no observation at the temperature (100°) most important for establishing it. It also appeared that there was a small difference at 0° c. between the lattice parameters of heated and unheated aluminium filings.

§ 2. MEASUREMENTS OF LATTICE PARAMETER

Further measurements were therefore made: they and the previous measurements are given in table 1. Both the old and the new measurements are divided

Table 1. The lattice parameter of aluminium

Temp. (° c.)	Lattice parameter					
	Unannealed filings		Annealed filings			
	OM	NM	OM	NM, OS	NM, NS	Mean, NM
0	$4{\cdot}038\ 9_1^{b}$	$4{\cdot}038\ 5_8^{C}$	$4{\cdot}039\ 1_5^{d}$	$4{\cdot}039\ 1_5^{a}$	$4{\cdot}039\ 1_2^{C}$	$4{\cdot}039\ 1_5$
50	—	—	—	$4{\cdot}043\ 7_0$	—	$4{\cdot}043\ 7_0$
100	$4{\cdot}048\ 5_5$	$4{\cdot}048\ 3_3$	—	$4{\cdot}048\ 5_0$	$4{\cdot}048\ 5_7$	$4{\cdot}048\ 5_4$
150	$4{\cdot}053\ 8_2$	$4{\cdot}053\ 5_3$	—	$4{\cdot}053\ 4_3$	$4{\cdot}053\ 3_0$	$4{\cdot}053\ 3_6$
200	$4{\cdot}058\ 9_9$	$4{\cdot}058\ 7_5$	$4{\cdot}058\ 9_9$	$4{\cdot}058\ 6_5$	$4{\cdot}058\ 5_5$	$4{\cdot}058\ 6_0$
300	$4{\cdot}070\ 1_9$	$4{\cdot}069\ 6_5$	$4{\cdot}069\ 9_4$	$4{\cdot}069\ 7_4$	$4{\cdot}069\ 7_3$	$4{\cdot}069\ 7_4$
400	$4{\cdot}081\ 9_7$	$4{\cdot}081\ 9_1$	$4{\cdot}081\ 9_9$	$4{\cdot}081\ 6_0$	$4{\cdot}081\ 6_7$	$4{\cdot}081\ 6_4$
500	$4{\cdot}094\ 6_8$	—				
600	$4{\cdot}108\ 6_3$	—	$4{\cdot}108\ 6_8^{e}$			
650	—	—	$4{\cdot}116\ 1_8^{c}$			

OM = Old measurements
NM = New measurements
OS = Old specimen
NS = New specimen
C = Cold-rolled sheet

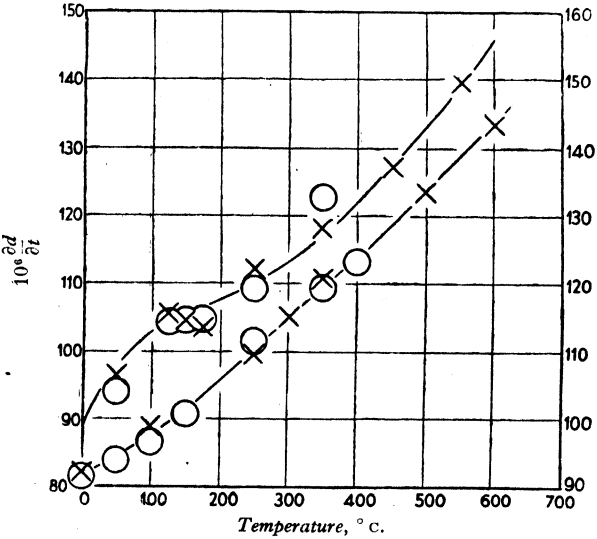
a = Mean of two
b = Mean of three
c = Mean of four
d = Mean of five
e = Mean of six

into two groups: those made before and those made after heating above 600° c. The experimental technique for the new measurements was as described by Stokes and Wilson (1941). The thermocouple was calibrated at a further fixed point: the melting point of zinc (419·45° c.).

It is rather disappointing to find that the new results from 100° upwards are, on the average, 0·000 3 A. lower than the old. In the previous paper it was estimated that the total error, both from the measurement of temperature and from the x-ray technique, was not greater than 0·000 2 A., but it appears that this was over-optimistic. The new measurements are probably the more reliable as the thermocouple was calibrated at a greater number of points.

§ 3. THE THERMAL EXPANSION COEFFICIENT

The mean values of $10^6\,\partial d/\partial t$ over intervals of 50° or 100° from both the old (×) and the new (○) measurements on unannealed aluminium filings are plotted in the figure (upper curve: scale on left). Both show a double curvature with a point of inflection at approximately 150° c. This is in general agreement with the curve published previously (full line). The values obtained from both the old (×) and the new (○) measurements on filings after heating to 600° are shown in the lower curve (scale on right) of the same figure. In addition to the means, values of $10^6\,\partial d/\partial t$ at the ends of the ranges are also plotted. These were calculated by the method of Bickley and Miller (1942), the first three differences being used. There is no sign of the double curvature shown by the upper curve. In view of the above-mentioned difference between the old and the new measurements, a difference in expansion between annealed and unannealed filings cannot be regarded as certain: it is, however, suggestive that both sets of measurements indicate a double curvature in the expansion coefficient of unannealed filings.



Slope of spacing—temperature curve. Upper curve, left scale, unannealed filings. Lower curve, right scale, annealed filings. ×, old measurements; ○, new measurements.

The coefficient of expansion, α, was obtained by dividing values of $\partial d/\partial t$ from the lower curve in the figure by the corresponding values of d. It is given in the first line of table 2. In the succeeding lines, values of $(1/l)(\partial l/\partial t)$ from the work of Hidnert (1925), Bollenrath * (1934), Taylor *et al.* (1938), Esser and Eusterbrock (1940–41) and Nix and MacNair (1941) are given. The present results are in excellent agreement with those of Hidnert and of Taylor *et al.*, the largest deviation of any of the three determinations from the mean being less than 2 %. Nix and MacNair's results agree well up to 300°, but their theoretical curve, based on a modification of Grüneisen's law, is low at higher temperatures. Its curvature is in the sense opposite to that observed. Bollenrath's results * are low above 100° and Esser and Eusterbrock's above 200°.

Table 2. The thermal expansion coefficient of aluminium, $\times 10^6$

Author(s)	Temp. 0°	100°	200°	300°	400°	500°	600°	% Al
Wilson	22·8	24·0	26·0	28·3	30·3	32·7	35·0	99·99
Hidnert	22·6	24·5	26·4	28·3	30·3	32·2	34·1	99·95
Bollenrath	22·8	23·6	25·0	26·6	28·9	31·5	—	99·87
Taylor *et al.*	23·2	24·3	25·9	27·9	30·4	33·4	—	99·996
Esser & Eusterbrock	23·4	24·5	25·8	27·4	29·2	31·3	33·7	99·87
Nix & MacNair (read from graph)	22·9	25·0	26·4	27·7	29·0	(30·3)	(31·6)	99·997

§ 4. THE EFFECT OF COLD-WORK ON THE LATTICE-PARAMETER

It will be noticed that the mean of the spacings of unannealed filings at 0° c. ($4{\cdot}038\ 9_1$ A.) is 0·000 2 A. lower than the mean of the spacings of annealed filings ($4{\cdot}039\ 1_5$ A.). The mean deviation is in each case 0·000 05 A. By any of the ordinary statistical tests the difference is significant, but in view of the difference between the old and the new results at higher temperatures it might be thought unwise to attach great importance to it. To produce a more definite effect, some of the aluminium was cold-rolled and the spacing of a small sliver measured. The result was $4{\cdot}038\ 5_3$ A. After heating to 220° it increased to $4{\cdot}039\ 1_2$ A. It would appear, therefore, that cold-work does produce an appreciable effect on the lattice parameter of aluminium. The Debye-Scherrer lines on the initial photograph of the cold-rolled sheet were slightly broader than those on the photograph obtained after heating. The broadening would be explained by a reduction of the crystallite size to about 1500 A., which is the order of magnitude found by Wood (1939) for other metals, but there are, of course, other explanations. A variation of the lattice parameter over a range of $\pm 0{\cdot}001$ A. would give about the same effect. Such a variation would be consistent with the observations of Wood and Smith (1941).

---

* It has been assumed that Bollenrath's table gives $(1/l)(\partial l/\partial t)$. It is possible, however, that it is $(1/l_0)(\partial l/\partial t)$.

REFERENCES

Bickley, W. G. and Miller, J. C. P., 1942. *Phil. Mag.* **33,** 1.
Bollenrath, F., 1934. *Z. Metallkunde,* **26,** 62.
Esser, H. and Eusterbrock, H., 1940–41. *Arch. Eisenhuttenw.* **14,** 341 (quoted in the *Metallurgist,* 27 June 1941).
Hidnert, P., 1925. *Sci. Pap. Bur. Stand., Wash.,* no. 497, 697.
Nix, F. C. and MacNair, D., 1941. *Phys. Rev.* **60,** 597.
Stokes, A. R. and Wilson, A. J. C., 1941. *Proc. Phys. Soc.* **53,** 658.
Tammann, G., 1934. *Z. Metallkunde,* **26,** 97.
Taylor, C. S., Willey, L. A., Smith, D. W. and Edwards, J. D., 1938. *Metals and Alloys,* **9,** 189.
Thomassen, L. and Wilson, J. E., 1933. *Phys. Rev.* **43,** 763.
Wilson, A. J. C., 1941. *Proc. Phys. Soc.* **53,** 235.
Wood, W. A., 1939. *Proc. Roy. Soc.* A, **172,** 231.
Wood, W. A. and Smith, S. L., 1941. *J. Inst. Met.* **67,** 315.

THE THEORETICAL VALUES OF THE PHYSICAL CONSTANTS

By Sir ARTHUR EDDINGTON, O.M., F.R.S.,
The Observatory, Cambridge

MS. received 8 *June* 1942

ABSTRACT. The principal physical constants are calculated by the theory developed in *Relativity Theory of Protons and Electrons* and subsequent papers, and are compared with the values given by Birge (1942) in *Reports on Progress in Physics,* **8.** There is satisfactory agreement with Birge's values as they stand. But the theory also indicates that small corrections are required in the computation of certain constants from experimental data ; when these corrections are included, the small discordance between the spectroscopic and deflection values of $e/m_e c$ disappears, and the agreement of observation and theory is complete. It is concluded that there ought to be no difference between the direct and indirect values of h/e. The calculated constant of gravitation is $6{\cdot}6665.10^{-8}$; and it is pointed out that the expected agreement of the calculated and observed values is not affected by the mean chemical constitution of the universe or the free radiant energy in space—a point previously left doubtful. As the purpose is not to justify or explain the theory but to compare it with observation, theoretical explanation is limited to the points which arise in adapting the theory to practical comparisons.

§ 1. INTRODUCTION

From the theory developed in *Relativity Theory of Protons and Electrons* (hereafter quoted as *P. & E.*) and subsequent papers (Eddington, 1937, 1940, 1940 a) it should be possible to determine precise values of the general physical constants. The recent publication of R. T. Birge's comprehensive survey of the experimental data (Birge, 1941) provides a suitable opportunity for tabulating the results and comparing them with observation.

A complication arises because the experimental values are expressed in three traditional units—the centimetre, gram and second—which have no relation to any theory. It is therefore necessary to select three measured quantities which, taken together, determine the centimetre, gram and second, to be used as " conversion constants ". They must be quantities which occur in the most fundamental part of physics, where the theoretical development is believed to be exact (a condition which excludes the characteristics of elements other than hydrogen); subject to this, they should be the most accurately ascertained constants available. Happily there is no doubt as to the most suitable choice. The three quantities are the velocity of light ($\mathbf{c}$), the Rydberg constant for hydrogen ($\mathfrak{R}$) and the Faraday constant for hydrogen ($\mathfrak{F}$). The last-named is the ordinary Faraday constant divided by the atomic weight of hydrogen, so as to reduce it to the scale $^1\mathrm{H}=1$ instead of $^{16}\mathrm{O}=16$. Birge gives—

Table 1

$\mathbf{c}=(2{\cdot}99776\pm{\cdot}00004)\,.\,10^{10}$ cm. sec.$^{-1}$

$\mathfrak{R}=109677{\cdot}58\pm{\cdot}05$ cm.$^{-1}$

$\mathfrak{F}=9573{\cdot}56\pm1{\cdot}0$ e.m.u.g.$^{-1}$

From table 1 it would be possible to reconstruct the C.G.S. system if the original standards were destroyed. We accordingly adopt it as a *verbal definition* of C.G.S. units which is in as close agreement as possible with the primary *gesticulatory definition.**

The calculated values of the physical constants, listed later, involve no observational data other than the three conversion constants. The uncertainty of any constant is therefore immediately evident from its dimensions in $\mathbf{c}$-$\mathfrak{R}$-$\mathfrak{F}$ units. The dimensions of the principal quantities are found to be

length	$\mathfrak{R}^{-1}$	angular momentum	$\mathfrak{R}^{-2}\,\mathfrak{F}^{-2}\mathbf{c}$
time	$\mathfrak{R}^{-1}\mathbf{c}^{-1}$	electrostatic charge	$\mathfrak{R}^{-1}\,\mathfrak{F}^{-1}\mathbf{c}$
mass	$\mathfrak{R}^{-1}\,\mathfrak{F}^{-2}$	constant of gravitation	$\mathfrak{F}^{2}\mathbf{c}^{2}$

If Birge's results are accepted, the uncertainties of $\mathbf{c}$ and $\mathfrak{R}$ are negligible, and $\mathfrak{F}$ has a probable error of 1 part in 10 000. The highest power of $\mathfrak{F}$ in any of the listed constants is $\mathfrak{F}^2$; their probable errors should therefore not exceed 1 part in 5000.†

A more serious complication arises because the so-called "observational values" of the constants are not derived wholly from observation, but are a combination of observational data with a theory which I regard as superseded. Thus a straightforward comparison with Birge's constants would not be an observational test of my theory, but a comparison partly with observation and partly with the theory which it condemns. It has to be recognized that the current procedure of computing constants from observational data has acquired the status of a vested interest: and, as far as practicable, we accept it as furnishing the official definition of the constants. We therefore go out of our way to adapt the theoretical formulae (§ 2) to the official definition—so that the calculated constants represent the results which Birge presumably would have obtained by his procedure of reduction if the observational data had been perfect. But the current procedure is not entirely self-consistent, and in the most accurate comparisons we cannot avoid applying small corrections to certain of Birge's values in order to obtain a genuine comparison of theory and observation. This inconsistency of the official definitions is perhaps rather fortunate, since it relieves us of the obligation of carrying respect for tradition to an embarrassing extreme.

---

* By a "gesticulatory definition" I mean pointing to an object and saying that the metre (kilogram, day) is the extension (inertia, period) of *that* object.

† Except the last three constants in table 3, for which the theory has not yet been extended to this degree of refinement.

§ 2. THE THEORETICAL FORMULAE

The formulae used in calculating the principal constants are—

$$(1)\quad e^2=\frac{1}{137}\frac{h\mathbf{c}}{2\pi},\qquad (2)\quad \mathrm{M}=\frac{136}{10}m_0,\qquad (3)\quad \mu'=\frac{1}{136}m_0,$$

$$(4)\quad \mathrm{M}=m_p+m_e,\qquad (5)\quad \mu=m_pm_e/(m_p+m_e),$$

$$(6)\quad \mu=\beta^{\frac{1}{4}}\mu',\qquad (7)\quad \beta=\frac{137}{136},$$

$$(8)\quad m_0=\frac{\frac{3}{4}\beta^{\frac{1}{4}}h\sqrt{(\frac{3}{5}N)}}{2\pi R_0\mathbf{c}},\qquad (9)\quad \frac{N_0}{R_0}=\frac{\pi\mathbf{c}^2}{\kappa\mathrm{M}}:$$

$$(10)\quad N_0=\tfrac{3}{4}N,\qquad (11)\quad N=2\times 136\times 2^{256},$$

$$(12)\quad \kappa=\frac{136\,.\,137}{10}\,\frac{\pi\mathfrak{F}^2\mathbf{c}^2\beta^{\frac{1}{4}}}{\sqrt{(\frac{5}{3}N)}},$$

$$(13)\quad \mathfrak{F}=e/\mathrm{M}\mathbf{c},\qquad (14)\quad \mathfrak{R}=\frac{1}{2}\left(\frac{1}{137}\right)^2\frac{\mu\mathbf{c}}{h}.$$

Most of the notation is defined in table 3. Of the remaining symbols, $\mathbf{c}$, $\mathfrak{R}$, $\mathfrak{F}$ have already been defined: and m_0 (the mass of a comparison particle), N (the cosmical number) and μ' may here be regarded merely as auxiliaries in the calculation.

Formulae (1), (2), (3), (4), (5), (8), (9) are *P. & E.* (15.65), (12.63), (12.62), (14.71), (14.72) with certain changes explained below; formulae (6), (7), (10) relate to these changes. The result (11) is given in *P. & E.*, §§ 16.5, 16.7. Formula (12) is redundant, being derivable from the others, but is given here as the most direct expression for the constant of gravitation κ in terms of atomic constants. Formula (13) follows from the definition of $\mathfrak{F}$, since M is the mass of a hydrogen atom: and (14) is the ordinary expression for the Rydberg constant simplified by using (1).

From formulae (2) – (6) we find that m_p, m_e are the roots of the equation

$$10m^2-136mm_0+\beta^{\frac{1}{4}}m_0{}^2=0. \qquad \ldots\ldots(15)$$

Apart from the β factors, (12) and (15) are *P. & E.* (14.75) and (12.47).

The changes that have been made in the *P. & E.* formulae are: firstly, a factor $\frac{3}{4}$ has been inserted in (8) and (10). This emendation was announced tentatively by Eddington and Thaxton (1940). I have since been able to simplify the theory, and it has become evident that the factor was wrongly omitted in my earlier work. Both in (8) and (10) it is introduced in reducing the spherical space of relativity theory to the flat space postulated in quantum theory. The number N_0 of elementary particles is now $\frac{3}{4}$ of the cosmical number N, which stands for the number of independent quadruple wave functions at any point. These wave functions are non-integrable in spherical space: and it is the non-integrability that causes the break-down of the simple exclusion principle which assigns one particle to each wave function.

Secondly, coefficients depending on the *Bond factor* $\beta = 137/136$ appear in (6) and (8), and consequentially in (12) and (15). In principle β^{-1} has the same origin as the factor $\frac{3}{4}$, both being of the form $n/(n+1)$ and arising out of the extension of a probability distribution from n to $(n+1)$ dimensions. The additional variate is introduced, in the one case when we represent the inertial-gravitational field quantum-mechanically instead of by an extraneous concept (curvature), and in the other case when we represent the electric field quantum-mechanically by an interchange co-ordinate. The primary introduction of β has already been noticed in *P. & E.*, §§ 15.8, 15.9; but the formulae there given correspond to $\mu = \beta\mu'$. Here the same factor is distributed, $\beta^{\frac{3}{4}}$ being included in (6) and $\beta^{\frac{1}{4}}$ in (8). This change is not an amendment of the theory, but an adaptation of it to give the officially defined constants instead of the constants which the theory itself points out as physically significant. It would be unsuitable in this paper to enter deeply into theoretical questions: but, since the β factors especially concern the comparison of theory with observation, we shall in § 6 go far enough to show how the factors $\beta^{\frac{1}{4}}$, $\beta^{\frac{3}{4}}$ have been determined.

The calculation of the constants proceeds as follows:—From (2), (3), (6) we obtain M/μ. Combining this with (14) we obtain $h/\mathrm{M}\mathbf{c}$. Since (13) gives $e/\mathrm{M}\mathbf{c}$, we obtain h/e: and by (1) we can then find h and e separately. The remaining steps are straightforward.

From R_0 we derive an interesting constant V_0, which is the limiting speed of recession of the galaxies. The well-known formula is

$$V_0 = \mathbf{c}/R_0\sqrt{3} \text{ cm. sec.}^{-1} \text{ per cm. distance.} \qquad \ldots\ldots(16)$$

In table 3, V_0 is given in the usual astronomical measure, km. sec.$^{-1}$ per megaparsec. Thanks to H. S. Jones's recent determination of the solar parallax, the result remains accurate to 1 part in 10 000. The observed speed is (as it should be) rather less than the calculated limiting speed; but the most optimistic estimate cannot put the probable error at less than 5 per cent.

The theory (Eddington, 1937) also gives two constants important in nuclear physics, namely the constants k and A in the expression

$$E = -Ae^{-r^2/k^2} \qquad \ldots\ldots(17)$$

for the non-Coulombian energy of two protons distant r apart. The formulae are

$$k = \frac{R_0}{\sqrt{N_0}}. \qquad A = \left(\frac{4}{3}\right)^{\frac{1}{2}} \frac{4}{\pi^{\frac{1}{2}}} \frac{m_p}{m_0} \frac{e^2}{k}. \qquad \ldots\ldots(18)$$

In the original investigation, A was given without the $(\frac{4}{3})^{\frac{1}{2}}$ factor, which has the same origin as the $\frac{3}{4}$ factor in (8) and (10). The β factor, if any, has not yet been considered: it can scarcely exceed $\beta^{\frac{1}{2}}$ in k, but may be more serious in A. Provisionally the uncertainties of k and A may be taken as 1 in 800 and 1 in 100, respectively. In the course of a general revision and simplification of the theory I have not yet dealt with A, and it is therefore more provisional than the other constants.

For completeness it may be added that the theory (Eddington, 1940 a) gives for the mass of a neutron $\mathrm{M} + \frac{3}{2}\mu$, or very nearly $m_p + \frac{5}{2}m_e$. More tentatively it gives 173·8 m_e for the mass of a mesotron, and 2·36 m_p for the mass of a "heavy mesotron" which decays into a proton or negatron. These results are not included in table 3.

§ 3. COMPARISON OF THEORY AND OBSERVATION

Setting aside the cosmological constants (R_0, N_0), the equations (1) to (14) boil down into a theoretical determination of three independent ratios $h\mathbf{c}/2\pi e^2$, m_p/m_e, $\kappa/\mathfrak{F}^2\mathbf{c}^2$. These, being purely numerical, do not involve the conversion constants, and therefore afford the most direct comparison of theory with observation. These are given in table 2 and compared with the results given by Birge, or obtained directly from his values of the constants.

The agreement seems not quite satisfactory, since the residuals of A and B are nearly twice their probable errors. Two such residuals out of three would be counted a significant discordance if they were independent. But the observed values of A and B both involve the observed constant $e/m_e\mathbf{c}$; and we find on examination that both residuals have the same source, namely a slightly discordant value of this constant. This is made evident by forming the combination $AB^{\frac{1}{3}}$, the observational value of which is independent of $e/m_e\mathbf{c}$.* The low probable error (1 part in 10 000) makes the test very stringent, and the agreement is perfect.

Table 2

Ratio	Calculated	Observed
$A=h\mathbf{c}/2\pi e^2$	137·000	137·030 ± ·016
$B=m_p/m_e$	1837·46	1836·56 ± ·56
$C=\kappa/\mathfrak{F}^2\mathbf{c}^2$	$8{\cdot}0939.10^{-37}$	$(8{\cdot}0981 \pm {\cdot}0060)10^{-37}$
$AB^{\frac{1}{3}}$	1678·01	1678.10 ± ·15

When the correlation of the residuals is allowed for, the agreement in table 2 is just about as close as the theory of probability admits.* But although the comparison does not indicate any real need for correcting the observed values, the theory insists that certain small corrections should be applied.

Birge's adopted value of $e/m_e\mathbf{c}$ is the mean of a deflection value and a spectroscopic value which are not very accordant. We shall find (§ 7) that the theory predicts a discordance of the two values with one another, and also a discordance with the value implied by the official definition of the other constants. Distinguishing the deflection and spectroscopic results (uncorrected) by the suffixes d and s, the theory in § 7 gives

$$(e/m_e\mathbf{c})_d = \beta^{-\frac{1}{24}}\, e/m_e\mathbf{c}, \qquad (e/m_e\mathbf{c})_s = \beta^{-\frac{1}{12}}\, e/m_e\mathbf{c}. \qquad \ldots\ldots(19)$$

Birge's values of $e/m_e\mathbf{c}$ when corrected in accordance with (19) become

spectroscopic	$(1{\cdot}75987 \pm {\cdot}00028)10^7$,
deflection	$(1{\cdot}76013 \pm {\cdot}00024)10^7$.

---

* It involves $\mathfrak{R}$, $\mathfrak{F}$ and the Avogadro number.

† It is an even chance that the worst of three residuals is less than 1·87 times its probable error.

Adopting the mean, $e/m_e\mathbf{c} = 1{\cdot}76000 \,.\, 10^7$ instead of Birge's uncorrected mean (1·7592), we obtain for the ratios *A* and *B* in table 2 the revised comparison:

	Calculated	Observed
$h\mathbf{c}/2\pi e^2$	137·000	137·009
m_p/m_e	1837·46	1837·40

I do not venture to give the probable errors of the observed values, which require technical discussion; but, although they are not quite so small as the agreement suggests, they are considerably smaller than in table 2, owing to the removal of the disagreement between the spectroscopic and deflection methods.

§ 4. VALUES OF THE CONSTANTS

The calculated values of the most important constants are given in table 3. All except nos. 16, 17 are in C.G.S. units. The authority for the observed values is Birge's *Report*, except no. 17 (Hubble and Humason, 1931) and nos. 18, 20 (Thaxton, 1940). The probable error, given in the last column, is in units of the last decimal place in the preceding column. The observed values are taken directly from Birge and the other authorities without correction.

Table 3

Ref. no.	Constant	Description	Calculated	Observed	P.E.
1	e	elementary charge (e.s.u.)	$4{\cdot}80333 \,.\, 10^{-10}$	4·8025	10
2	h	Planck constant	$6{\cdot}62504 \,.\, 10^{-27}$	6·6242	24
3	M	mass of ^{1}H atom	$1{\cdot}67368 \,.\, 10^{-24}$	1·67339	31
4	μ	$m_p m_e/(m_p+m_e)$	$9{\cdot}09873 \,.\, 10^{-28}$	..	..
5	m_e	mass of electron	$9{\cdot}10368 \,.\, 10^{-28}$	9·1066	32
6	m_p	mass of proton	$1{\cdot}67277 \,.\, 10^{-24}$	1·67248	31
7	κ	constant of gravitation	$6{\cdot}66649 \,.\, 10^{-8}$	6·670	5
8	$h\mathbf{c}/2\pi e^2$	fine-structure constant	137	137·030	16
9	h/e	. . . .	$1{\cdot}37926 \,.\, 10^{-17}$	1·37933	23
10	m_p/m_e	mass-ratio	1837·46	1836·56	56
11	$e/m_e\mathbf{c}$	(deflection method)	$1{\cdot}76006 \,.\, 10^{7}$	1·75959	24
12	$e/m_e\mathbf{c}$	(spectroscopic method)	$1{\cdot}76006 \,.\, 10^{7}$	1·75880	28
13	N_0	particles in universe	$2{\cdot}36216 \,.\, 10^{79}$	..	..
14	M_0	mass of universe	$1{\cdot}97675 \,.\, 10^{55}$	..	..
15	R_0	Einstein radius of space	$9{\cdot}33544 \,.\, 10^{26}$	..	..
16	R_0	(in megaparsecs)	302·38	..	..
17	V_0	nebular speed (km. sec.$^{-1}$ mp.$^{-1}$)	572·36	560	..
18	k	nuclear range-constant	$1{\cdot}92079 \,.\, 10^{-13}$	1·9	..
19	A	nuclear energy-constant	$4{\cdot}2546 \,.\, 10^{-5}$	..	..
20	$A/m_e\mathbf{c}^2$	. . . .	52·006	51·4	..

The main purpose of table 3 is to give explicitly the calculated values. With the exception of nos. 8, 10, 13, they involve the conversion constants in table 1; but the effect of any future changes in the adopted values of $\mathfrak{F}$, $\mathfrak{R}$, $\mathbf{c}$ is readily found from the physical dimensions of the constant considered, as explained in § 1.

Nos. 1 to 12 are all obtained by combining the three ratios in table 2 with the three conversion constants. Thus the comparison with Birge's values adds nothing essentially new to the discussion in § 3: the same agreement and the same mild discordance with the uncorrected observed values is shown (with much redundancy) in another form.

Besides the accurate comparisons, the observational determinations of nebular recession and of the two nuclear constants provide independent but much rougher comparisons (nos. 17, 18, 20). It is worth remarking that by equations (16) and (18)

$$kV_0 = \mathbf{c}/\sqrt{(3N_0)}, \qquad (N_0 = \tfrac{3}{2}\times 136 \times 2^{256}), \qquad \ldots\ldots(20)$$

so that the range-constant of the proton-proton force can be found directly from the speed of recession of the galaxies, or vice versa, with no other observational data except the velocity of light.*

An improvement in the observational determination of κ is greatly to be desired. It has been suggested to me that the observed value should differ from my calculated value by an amount depending on the mean packing-fraction of the elements of which the universe is composed: but I find that the suggestion is fallacious. In the theoretical calculation the universe was represented as a uniform static zero-temperature distribution of protons and electrons. It is, of course, legitimate to rearrange the matter of the universe so as to simplify the theoretical determination, just as it is legitimate to rearrange the matter in the laboratory so as to simplify the observational determination: but in such rearrangement the experimenter cannot, and the theorist must not, violate the conservation of energy. There is therefore an implicit assumption that the energy of free radiation, cosmic rays, etc., in space is just sufficient to transmute the complex elements back into hydrogen, restore the gravitational energy lost in the formation of condensations, and leave a uniform distribution at zero temperature. For theoretical reasons I believe this to be true; but I should now describe it as a *consequence* of the formula for κ, and not an assumption that must be made in order to derive it. In formula (12) all reference to the " universe " is eliminated, and the ratio $\kappa/\mathfrak{F}^2\mathbf{c}^2$ of locally measured constants is expressed in terms of the number N of independent quadruple wave functions introduced by our system of analysis. The fact that the wave functions are non-integrable emphasizes that N (unlike N_0) is a local concept. We should ordinarily say that the remote environment has no effect on $\kappa/\mathfrak{F}^2\mathbf{c}^2$, because it has the same effect on $\mathfrak{F}^2\mathbf{c}^2$ as on κ.

§ 5. THE PROBLEM OF CONSISTENCY

Mass, momentum, energy, electric charge, etc., are primarily defined in molar physics. In some cases the molar definition can be adapted to cover microscopic quantities. Thus the mass M of a hydrogen atom is $1/n$ of the molar mass of a quantity of hydrogen composed of n atoms: and the mass of any *neutral* atom or particle can be defined similarly. The elementary charge e is $1/n$ of the molar electric charge of a body from which n electrons have been removed. Microscopic quantities defined in this way are said to be " molarly controlled ". There is one important microscopic constant which is not molarly controlled, namely

---

* For an accurate determination, a correction must be applied to the observed nebular recession to obtain the limiting speed of recession. Ideally the correction can be found by astronomical determinations of the average density of matter in space, but these are at present too inaccurate.

m_e. The mass of an electron is obviously not defined as $1/n$ of the molarly measured mass of an aggregation of n electrons: the observer (if he escaped electrocution) would have no idea how to set about so unprecedented a measurement.

In quantum theory a fresh start is made, and the terms mass, energy, charge, etc., are applied to *analogues* of the molar quantities, which play a corresponding part in the equations of quantum dynamics to that played by the originals in classical dynamics. Whether the analogy is to be construed as identity (for such quantities as are molarly controlled) can only be ascertained when quantum theory is extended into a unified microscopic and molar theory as in *P. & E.* We must not assume the identity prematurely.

There is usually a definite understanding that, whatever other divergences may exist, lengths and periods in microscopic structures are molarly controlled. This is implied when we assume that the frequencies of light waves emitted by an atom are equal to the frequencies of oscillating sources in its structure; for the undulatory theory of light is part of molar theory. This assumption or convention was the origin of Heisenberg's theory, which superseded Bohr's theory of the atom; and it is too deeply rooted in wave mechanics to be disturbed. Molar control of lengths and periods is imposed through the formula (14) for the Rydberg constant.

We naturally try to find a set of physical constants which, besides satisfying molar control of lengths and periods, shall (as a minimum) be consistent with

(1) the analogy between quantum dynamics and molar dynamics;
(2) molar control of M and other neutral masses;
(3) molar control of e;

but it is by no means clear that even these minimum conditions can be satisfied simultaneously. Accepting (1), the measure-scale of mass can certainly be adjusted to satisfy (2); but it appears that a rather different scale is required to satisfy (3). This is the cause of the corrections introduced in equation (19).

For the theorist, it is impossible to sacrifice (1), and evidently (3) is the condition to be dropped. This means that the Faraday constant $\mathfrak{F} = e/\text{M}c$ used in microscopic physics will not agree with the constant found by molar observation, and there will be corresponding differences in the constants e and h. This, however, is a small price to pay for the far-reaching developments of quantum mechanics which can only be utilized by accepting condition (1).

But, notwithstanding its advantages, I clearly cannot adopt this system of definition of the constants in the present context. To tamper with the Faraday constant would bar out all comparison with Birge's constants at the very outset. I must follow him in accepting conditions (2) and (3) as a *sine qua non*. This will lead to trouble sooner or later, when we try to utilize results predicted by quantum theory on the basis of condition (1). The trouble does not begin immediately, because we have a constant m_e which is not molarly controlled, and is (on this system) virtually defined so as to satisfy the most immediate requirements of condition (1). The formulae of § 2 and the calculated values in table 3 relate to this system of constants, which we shall call the " official system ". It will be treated in § 6. The " trouble " which arises in the adaptation to the official system of results which are based on condition (1) will be treated in § 7.

To sum up:—We accept the ruling that in microscopic physics lengths, periods and the masses of electrically neutral objects must agree with the molar definitions.* As regards electric charges and the masses of charged particles, we have a choice of two evils. The quantities which appear in the equations of quantum theory are the analogues (not the direct equivalents) of molar charge and mass, and the theorist would naturally prefer that this should be the recognized definition of constants such as m_e and e. But in the official system of constants, e is defined as the direct equivalent of molar charge. (There is no direct molar equivalent of m_e). Thus the results of quantum theory normally require a correction to reduce them to the official system of constants. Although the formula for $\mathfrak{R}$ is a result of quantum theory, it is in an exceptional position because of the adoption of $\mathfrak{R}$ as a conversion constant. This requires that the quantity determined by it, namely $\mu \mathbf{c}/h$, shall be the same in both systems of constants.

§ 6. THE β-COEFFICIENTS

The constant denoted by m_p/m_e is usually defined in the way indicated by the notation—a definition which would have had no meaning before the isolation of the electron. But the same constant can be described, independently of the atomic structure of electricity, as the ratio of the *densities* of the electrically positive and negative material in molar hydrogen. The related constant M/μ† can similarly be described as a density-ratio of external and internal energies in molar hydrogen. Relativity mechanics is based on the energy tensor, which includes the density as a component; so that in relativistic theory m_p/m_e and M/μ appear primarily as density-ratios, and their interpretation as mass-ratios comes later. By rearrangement of the theory in *P. & E.*, I now obtain equations (2) and (3), which yield $\text{M}/\mu' = 136^2/10$, *before* treating microscopic problems.

In passing over into microscopic physics, we introduce the extra variate referred to in § 2, and this determines the scale of microscopic structure. Like all observables, it has an uncertainty: and its introduction involves an extra dimension of the domain of probability-distribution of the complete energy tensor, extending the number of degrees of freedom from 136 to 137. We find that this alters the measure of the energy tensor, and therefore of the density, in the ratio β; so that the density μ' is replaced in microscopic theory by the density $\mu_0 = \beta\mu'$. This is a distinction between the analogue μ_0 and the equivalent μ' of the molar density.

My earlier work in *P. & E.* did not go beyond this point; but it is clear that when we modify the definition of density to suit microscopic theory some further adjustment will become necessary. Either the constant of gravitation must be modified, or the metric of space will be changed. By accepting the condition of molar control of e, we settle the way in which this adjustment is to be made. For e^2 is of dimensions ML^3T^{-2}; and in order that its measure may be unaltered, the change of measure of density (dimensions ML^{-3}) in the ratio β must be coupled with a change of L^6T^{-2} in the ratio β^{-1}. To preserve relativistic symmetry L and T must be altered in the same ratio; hence the increase of density in the ratio β is coupled with a decrease of lengths in the ratio $\beta^{-\frac{1}{4}}$. It may be noticed that this combination makes Action invariant—a highly satisfactory property from the ordinary dynamical point of view.

---

* This does not apply to constants or systems of measurement employed in auxiliary theoretical investigations (e.g. μ_0, $\mathfrak{R}'$ in § 6) which are eliminated in the final formulae.

† To six significant figures $\text{M}/\mu = m_p/m_e + 2$.

The constant $\mathfrak{R}$ in table 1 is found by a molar measurement of light in its molar aspect (undulatory theory). Being the reciprocal of a length, it is transformed into a constant $\mathfrak{R}'=\beta^{\frac{1}{4}}\mathfrak{R}$ when μ' is transformed into $\mu_0=\beta\mu'$. The relation found in the theory of the internal wave functions of a hydrogen atom is therefore primarily

$$\mathfrak{R}'=\frac{1}{2}\left(\frac{1}{137}\right)^2\frac{\mu_0\mathbf{c}}{h}, \qquad \ldots\ldots(21)$$

since h (which has the fixed relation (1) to e) is an invariant of the transformation. This gives

$$\mathfrak{R}=\frac{1}{2}\left(\frac{1}{137}\right)^2\frac{\beta^{\frac{3}{4}}\mu'\mathbf{c}}{h}. \qquad \ldots\ldots(22)$$

Comparing with (14), we have $\mu=\beta^{\frac{3}{4}}\mu'$, which is the equation (6) that has been used in our calculations. We notice that molar control of e (and h) is secured by the choice of transformation; molar control of period and length is imposed by making (22) agree with (14); and molar control of mass is in this problem irrelevant, since we are not dealing with a neutral particle.

Equation (8) is made comparable with (21) by writing it in the form

$$R_0^{-1}=\text{numerical constant}\times\frac{m_0\mathbf{c}}{h}.$$

The theory gives primarily a constant $R_0'^{-1}=\beta^{\frac{1}{4}}R_0^{-1}$, just as it gives $\mathfrak{R}'=\beta^{\frac{1}{4}}\mathfrak{R}$; so that the factor $\beta^{\frac{1}{4}}$ is introduced when we employ the ordinary Einstein radius R_0. There is no β factor in (9), which is a well known equation of molar relativity theory. The factor $\beta^{\frac{1}{4}}$ in (8) is carried through into (12), which is derived from (8), (9) and (13).*

This shows how the β-coefficients in the equations of § 2 have been determined, and we proceed to the explanation of the β-coefficients in equation (19).

§ 7. THE CONSTANT $e/m_e c$

In a comprehensive treatment we should begin with the complete set of theoretical relations expressed in terms of the constants which the theory itself selects, and then, considering each kind of experimental measurement in turn, investigate the precise expression of the result of the measurement in terms of those constants. The existing treatment is an attempt at a short cut. Out of the wide variety of experimental measurements, three or four are selected on account of the great practical accuracy attained: constants directly related to these selected measurements are adopted, and the theorist is expected to transform his formulae so that they are expressed in terms of those constants. This

---

[*] Since the observed value of κ is relatively inaccurate, it is specially important to control the foregoing calculation of the β-coefficient in (12). I may therefore add that the simplest calculation employs the system of constants in § 7. The coefficient is then $\beta^{\frac{1}{4}}$; but the corresponding charge is e_s, so that $\mathfrak{F}$ must be replaced by $\mathfrak{F}_s$. By (28) we have $\beta^{\frac{1}{4}}\,\mathfrak{F}_s^2=\beta^{\frac{1}{4}}\,\mathfrak{F}^2$, so that the formula is confirmed.

is a troublesome policy in the long run; for sooner or later one of the non-selected experimental measurements becomes the centre of attention. The special adaptation of the theory to the selected measurements becomes irrelevant, and we have to transform the formulae back to their original form before we can relate them to the new measurement. At present the kinds of measurement accurate enough to be used in determining the constants are very limited, and the trouble is localized: it is concerned with μ, and with m_e which is derived from μ.

A " spectroscopic " value of M/μ (or m_p/m_e) can be deduced from the ratio of the Rydberg constant for deuterium (or helium) to the Rydberg constant for hydrogen. Since this is a deduction of one purely numerical ratio from another, the theory is particularly simple; only we must use the primary theoretical formulae which satisfy condition (1) of § 5, and not the formulae adapted for purposes irrelevant to this particular measurement. The following remarks therefore relate to the primary theory.

To avoid unnecessary complication we employ " natural units " (Eddington, 1940) such that

$$\mathbf{c}=1, \qquad \kappa h^2=1. \qquad \ldots\ldots(23)$$

This leaves one disposable unit; and it is easily found that when the unit is changed so that densities are changed in the ratio β, masses are changed in the ratio $\beta^{\frac{1}{4}}$, and lengths in the ratio $\beta^{-\frac{1}{4}}$. Also h and e change in the ratios $\beta^{\frac{1}{4}}$, $\beta^{\frac{1}{8}}$. Proceeding as in § 6, we have in place of (22)

$$\mathfrak{R}=\frac{1}{2}\left(\frac{1}{137}\right)^2\frac{\mu'\mathbf{c}\beta^{\frac{1}{6}}}{h_s}, \qquad \ldots\ldots(24)$$

where h_s is the natural theoretical value of h in microscopic theory (the symbol h having already been appropriated to an adapted value). Comparing with (22), we have

$$h_s=\beta^{\frac{1}{12}}h. \qquad \ldots\ldots(25)$$

The difference is that h corresponds to conditions (2) and (3) in § 5, and h_s corresponds to conditions (1) and (2). Since e is always related to h by $h\mathbf{c}/2\pi e^2=137$, we have $e_s=\beta^{\frac{1}{24}}e$. The value e is the equivalent of molar charge (i.e. it is molarly controlled), and the value e_s is the analogue of molar charge employed in quantum dynamics. The factor $\beta^{\frac{1}{12}}$ in (25) is the result of taking out the transformation of length in the ratio $\beta^{\frac{1}{4}}$ imposed by condition (3) before putting in the transformation in the ratio $\beta^{\frac{1}{6}}$ imposed by condition (1). Since condition (2) is common to the two systems, M is unchanged; and μ' is also unchanged, since the relation $\mu'=10\mathrm{M}/136^2$ is independent of microscopic theory as explained at the beginning of § 6.

The " observed " spectroscopic value of M/μ or m_p/m_c is computed on the assumption that the classical formulae (4) and (5) for transforming a two-particle system into external and internal particles apply. This assumption commits us to condition (1); for evidently a classical formula will not apply to quantities in quantum theory unless they are the analogues of the classical quantities. On this understanding it is easily proved that the assumption is valid.* Thus the uncorrected spectroscopic constants $(\text{M}/\mu)_s$ and $(m_p/m_e)_s$ correspond to the values h_s, e_s, and are inconsistent with the system of constants in table 3 which corresponds to $\hbar$ and e. In particular, the corresponding Faraday constant $\mathfrak{F}_s = e_s/\text{M}\mathbf{c}$ is $\beta^{\frac{1}{24}}\mathfrak{F}$.

By (13), (14) and (1)

$$\mathfrak{F}^2\mathfrak{R} = \frac{1}{2}\left(\frac{1}{137}\right)^2 \frac{e^2}{h\mathbf{c}}\frac{\mu}{\text{M}^2} = \frac{1}{4\pi}\left(\frac{1}{137}\right)^3 \frac{\mu}{\text{M}^2}. \qquad \ldots\ldots(26)$$

Since $(\text{M}/\mu)_s$ corresponds to $\mathfrak{F}_s$ instead of $\mathfrak{F}$, and the other constants are unchanged, we have

$$\frac{(\text{M}/\mu)_s}{\text{M}/\mu} = \frac{\mathfrak{F}^2}{\mathfrak{F}_s^{\,2}} = \beta^{-\frac{1}{12}}. \qquad \ldots\ldots(27)$$

and $(m_p/m_e)_s$ differs from m_p/m_e in nearly the same ratio. The collected comparison is

$$h_s = \beta^{\frac{1}{12}} h, \qquad \mu_s = \beta^{\frac{1}{12}} \mu, \qquad e_s = \beta^{\frac{1}{24}} e, \qquad \mathfrak{F}_s = \beta^{\frac{1}{24}} \mathfrak{F}. \qquad \ldots\ldots(28)$$

Applying these factors to the constants already calculated, we obtain

Table 4

$\mathfrak{F}_s = 9576{\cdot}48$	
$e_s = 4{\cdot}80480 \,.\, 10^{-10}$	
$h_s = 6{\cdot}62908 \,.\, 10^{-27}$	
$(m_p/m_e)_s = 1836{\cdot}34$	
$(e/m_e\mathbf{c})_s = 1{\cdot}75899 \,.\, 10^7$	(Birge $1{\cdot}75880 \pm {\cdot}00028$)
$(e/m_e\mathbf{c})_{ss} = 1{\cdot}75953 \,.\, 10^7$	(Birge $1{\cdot}75959 \pm {\cdot}00024$)

The last two quantities require explanation. Birge does not give the observational value of $(m_p/m_e)_s$ explicitly, but combines it with the ordinary Faraday constant so as to obtain "the spectroscopic value of $e/m_e\mathbf{c}$". This is a preposterously hybrid constant, since the e of the Faraday constant and the m_e of the spectroscopic result belong to different systems. However, we can calculate it for comparison with Birge. The result, denoted by $(e/m_e\mathbf{c})_s$, agrees excellently.

The *pure* spectroscopic value, which we denote by $(e/m_e\mathbf{c})_{ss}$, is obtained by combining $(m_p/m_e)_s$ with $\mathfrak{F}_s$. This is the value which we should expect to obtain by a measurement such as the deflection experiment which determines $e/m_e\mathbf{c}$ directly. This conclusion requires careful verification, since the comparison of a dimensional quantity with observed measurements is a more complicated matter than the comparison of a numerical ratio m_p/m_e. The scrutiny is as follows :—

The quantities actually measured in the deflection experiment are molar electromagnetic fields X, H and the curvature of track of a wave-packet. The curvature measurement applies directly to the theoretical equations, since lengths in those equations are molarly controlled. But it might be questioned whether X and H can be used as they stand, since we have seen that $\mathfrak{F}$ is inapplicable. Consider a modification of the experiment in which the fields are produced by charged bodies whose positions, velocities and molar charges ne are measured, so that n is determined from the molar measurement of charge by using the known constant e. The positions, velocities and n are transferred directly into the microscopic equations, and we can deduce from the experiment the pure spectroscopic constant $(e/m_e\mathbf{c})_{ss}$. It is clear that the two forms of experiment must give the same result, for a discrepancy would imply that it makes a difference whether we measure X and H or measure the molar charge and velocity which produce them; that is to say, it would be a discrepancy in Maxwell's theory. Thus X and H require no correction, the adjustment to microscopic theory being automatically provided for in the fact that they are computed (or measured) as though the body producing them had a charge ne instead of ne_s.

---

* Valid in so far as the only relevant difference between the proton and the deuteron is the difference of mass. I do not profess to have examined whether the effect of the complexity of the nucleus has been rigorously allowed for in the reductions.

Since $(e/m_e\mathbf{c})_{ss} = (e/m_e\mathbf{c})_d$, we have inserted Birge's value of $(e/m_e\mathbf{c})_d$ for comparison in table 4. In § 3 the comparison was made the reverse way by applying corrections to Birge's values of $(e/m_e\mathbf{c})_s$ and $(e/m_e\mathbf{c})_d$ to obtain the official constant $e/m_e\mathbf{c}$.*

In table 2 (p. 120) of Birge's *Report*, twelve determinations of $e/m_e\mathbf{c}$ by seven distinct methods are collected, and the table conveniently indicates how the results depend on the Faraday constant. We have not scrutinized each method separately, but there seems no reason to suspect any theoretical error except the use of an inconsistent Faraday constant. A constant larger in the ratio $\mathfrak{F}_s/\mathfrak{F}$ or $\beta^{\frac{1}{24}}$ should have been employed. This correction should make all the determinations agree. The amended values correspond to $(e/m_e c)_d$, and a further multiplication by $\beta^{\frac{1}{24}}$ is necessary to reduce them to official reckoning.

§ 8. THE CONSTANT *h/e*

Another question of practical importance is whether the " direct value " of h/e should agree with the indirect value. According to Birge, the latest experiments give no indication of discrepancy; but in view of the stop-press character of this information, and the many vicissitudes of opinion, we desire an independent theoretical answer.

I assume that "direct" refers to a determination $h/e = \delta V/\mathbf{c}\delta\lambda^{-1}$, where $\delta\lambda^{-1}$ is the measured change of spectral wave-number due to the fall of an electron through a measured potential difference δV. We remark first that the equation does not introduce the constants μ, m_e, which are the root of the divergence of the theoretical and official systems. The theory of the experiment is in fact independent of the quantum dynamics of the atom, which only appears as a catalyst in the conversion of a molarly measured energy eV into a molarly measured energy $h\nu$. Also, if the experiment could be satisfactorily carried out with hydrogen, the whole theory of it would fall within the part of quantum theory where the official constants directly apply. In any of these ways we can see that the direct value will agree with the official indirect value in table 3. Alternatively we can eliminate h by the fixed relation $hc/2\pi e^2=137$, and regard the experiment as a determination of e by the equation

---

* The formulae (19) are not quite exact, owing to the difference of transformation of M/μ and m_p/m_e; but the inaccuracy does not affect the first six significant figures.

$$e=\frac{1}{137}\frac{\delta V}{2\pi\delta\lambda^{-1}},$$

and it is clear that e will be in the same molar electrical measure in which V is expressed.

The theory therefore predicts agreement of the direct and indirect values.

REFERENCES

Birge, 1941. *Rep. Progr. Phys.* **8**, 90.
Eddington, 1937. *Proc. Roy. Soc.* A, **162**, 155.
Eddington, 1940. *Proc. Roy. Soc.* A, **174**, 16.
Eddington, 1940 a. *Proc. Roy. Soc.* A, **174**, 41.
Eddington and Thaxton, 1940. *Physica*, **7**, 122.
Hubble and Humason, 1931. *Astrophys. J.* **74**, 76.

NOTE ON SOME RECENT PAPERS ON PHYSICAL QUANTITIES AND THEIR DIMENSIONS

By W. E. DUNCANSON, M.Sc., Ph.D.,
University College, London (at University College of North Wales, Bangor)

MS. received 14 *May* 1942

ABSTRACT. Examination of recent papers on dimensions and physical quantities suggests a need for more careful definition of these quantities. This matter was dealt with in an earlier paper, but is now expanded and treated in more detail. It is maintained that if definitions were more precise there would not be the same confusion about the dimensions of certain quantities. There may still be differences of opinion on the most appropriate definition of a particular quantity, but if the definition is adequate then there is no doubt about the dimensions of the quantity as deduced from the definition. In this connection there is discussed the use of quantities defined as ratios, such as specific gravity (or relative density), and their more fundamental equivalents, such as density. Further, a logical set of definitions leads to the conclusion that the introduction of certain universal constants is unnecessary—for example, the Newtonian constant N in the equation for force, $F=Nmf$. Finally, there is discussed briefly the definition of the intensive quantity, temperature, in terms of the extensive quantity, energy.

§ 1. INTRODUCTION

From consideration of the recent papers on dimensions and on the nature of physical quantities*, there are two points which emerge:

(*a*) That unnecessary quantities are introduced into physics, partly because of lack or ambiguity of definition, and

(*b*) That much of the discussion on the dimensions of physical quantities is misplaced in that the difference of opinion on dimensions derives from differences in the definition of the quantities, and the emphasis of much of the discussion which takes place on dimensions should be moved a step further back to the definitions.

---

* Brown (1941), Yarnold (1942), Benham (1942) and Hansel (1942).

These two points are not entirely independent, but, to begin with, it will be convenient to treat the first apart from any dimensional considerations.

2. UNNECESSARY QUANTITIES IN PHYSICS

Let us consider the following four expressions

$$F=G\frac{M_1M_2}{d^2}, \qquad \ldots\ldots(1)$$

$$F=NMf, \qquad \ldots\ldots(2)$$

$$F=\frac{a}{K}\frac{q_1q_2}{d^2}, \qquad \ldots\ldots(3)$$

$$F=\frac{b}{\mu}\frac{m_1m_2}{d^2}, \qquad \ldots\ldots(4)$$

where M, q, m, f, F are respectively mass, charge, magnetic pole, acceleration, force, and G, N, a, b are constants.

Benham (1942) insists on the necessity for retaining both G and N, while Yarnold (1942) requires that a and b should appear in the appropriate equations.

First let us consider equations (1) and (2). Assuming mass, length and time as indefinables, there are three quantities remaining which require definition, viz. F, G and N. As was pointed out in a previous paper (Duncanson, 1941), one equation (such as, for instance, equation (1)) cannot be used to define two quantities. Any equation must either express a relationship between known quantities, or define one quantity in terms of the other *known* quantities in the equation—or, more generally, the number of independent equations must equal the number of quantities to be defined. Now, unless F is otherwise defined, equation (1) contains two undefined quantities, F and G. The only other way of defining F is by using a method such as that expressed in equation (2) (or its equivalent in terms of momentum). If, then, we define force in terms of mass and acceleration, why is it necessary to introduce a further quantity N to confuse the issue? If an equation such as (2) is used to define force, then *force equals the product of mass and acceleration* can be taken as the definition of force, and there is no question of further quantities involved. To press the point further, if N is introduced into (2), why should not density be written as

$$\rho=C\frac{M}{V},$$

where C is a physical quantity with dimensions?

If it is still insisted that G and N must both be retained, then a definition of force must be given which is independent of both of them. Hansel (1942), for instance, states (p. 168): "Newton's first law of motion provides a definition of mass and a definition of force", and then, assuming that force has been adequately defined, proceeds to state Newton's second law (p. 175): "That the ratio of the force acting on a body to the acceleration which that force gives to the body is constant". Now Newton's first law certainly does not define the physical quantity, force. It introduces us to the concept of force, but there is no definition of such a quantity in terms of other previously defined quantities, and, therefore, the second statement cannot be regarded as a law, but as a definition of force.

Let us now consider equation (3)—and remarks made about this equation will be equally applicable to (4). It is unnecessary to introduce both a and K into (3). Some quantity is needed to take care of the variation of force with the material of the dielectric, but Yarnold does not appear to justify the inclusion of a. If it were necessary to take into account some other variable quantity, then a would have to be retained. In this respect I am in agreement with Guggenheim (1942) when he suggests that two distinct quantities are, in practice, used in this connection, viz. permittivity, which we shall denote by k, and specific inductive capacity (or dielectric constant), denoted by K.

The logical procedure is to define the permittivity k by the equation $F = q_1 q_2 / kd^2$, and the specific inductive capacity K by $K = k/k_0$, where k_0 is the permittivity of a vacuum.

Similar pairs of quantities are used in the case of density and specific gravity, and of thermal capacity per unit mass and specific heat, where *specific* is used in each case to denote the ratio. It would be more satisfactory to use *relative* rather than *specific* to indicate the various ratios: for example, relative density, relative thermal capacity per unit mass, relative permittivity. But our chief concern here is not nomenclature but emphasis of the fact that two different quantities are often used for the same purpose.

This practice of expressing the magnitude of a physical quantity in terms of its magnitude for a particular substance under given conditions is quite common in physics, although it is not always as obvious as in the instances given immediately above. For example, the expression of resistance in terms of that of mercury under certain conditions; or again, the electric current in terms of the current necessary to deposit a certain mass of silver in a given time. These, together with specific inductive capacity, will be recognized as being on the same footing as specific gravity, which is, simply, the mass of a given volume of the substance in terms of the mass of the same volume of water.

The point it is desired to emphasize here is that these ratios must be recognized as distinct from the fundamental quantities concerned, and care should be taken when introducing them into equations, as the two types of quantity are not interchangeable.

§ 3. DIMENSIONS AND DEFINITIONS

Once we have defined our quantities with certainty, there is then no doubt about their dimensions, which brings us to our second point, namely, that any differences in the dimensions of a quantity result from differences in the definition, and that discussion should centre round the definition rather than the dimensions. A case in point is that of temperature. Benham (1942), Brown (1941) and Duncanson (1941) give $[L^{-1}]$, $[L^2T^{-2}]$, and $[ML^2T^{-2}]$ respectively for the dimensions of temperature. Three possibilities arise;

(*a*) That the respective definitions from which these dimensions are deduced define different quantities. The point at issue would then be which was the most desirable quantity to use from a practical and a theoretical point of view, a consideration we will not pursue further here.

(*b*) That some or all of the definitions are inadequate,

(*c*) That different dimensions are derived from equivalent definitions owing to the presence in the definition of quantities whose own dimensions are disputed. But this only throws the discussion a stage further back, to the definition of these quantities.

With these possibilities in mind, let us examine the three following definitions of temperature from which result the three sets of dimensions quoted above.

1. Brown defines temperature in the following manner. Firstly, specific heat is defined as a ratio, thus having no dimensions: then temperature is defined from the equation

$$\text{mass} \times \text{specific heat} \times \text{temperature} = \text{energy},$$

and from this the dimensions of temperature follow immediately as $[L^2T^{-2}]$.

2. The present writer does not believe that these two definitions can be used in conjunction with each other, and defines the temperature of a gas in terms of the mean energy of the molecules of that gas (actually 2/3 of the mean energy).* Hence it is seen that the dimensions of temperature are the same as those of energy, namely $[ML^2T^{-2}]$.

3. Benham introduces temperature with the statement: "We shall now take temperature as energy per unit mass, understandable from the kinetic theory and from calorimetry with heat as a form of energy". From this he deduces that the dimensions are those of $\left[\frac{\text{energy}}{\text{mass}}\right]$, which appears to be the same as those deduced by Brown, but on closer examination there emerge several distinct differences. Firstly, by the introduction of the Newtonian constant N, Benham deduces the dimensions of energy as $[ML^{-2}]$, and hence those of temperature as $[L^{-2}]$: this difference arises, not primarily from the definition of temperature, but from the dimensions ascribed to energy. Secondly, his statement of temperature as it stands is not at all clear. Assuming the normal concept of temperature is intended, then the statement is not true if it is meant that the temperature of a substance is the energy per unit mass of that substance; two different substances in thermal contact would then be at different temperatures. The alternatives are that Brown's definition is intended, or that the temperature is to be defined in terms of the energy per gram molecule of the substance—consistent with the dimensions of $\left[\frac{\text{energy}}{\text{mass}}\right]$ if, as is often assumed, gram molecule has the dimensions of [mass]. But it has already been pointed out in our previous paper that gram molecule is often used as a conventional term for denoting a given number of molecules, and, therefore, when used in this sense it is a pure number and has no dimensions. After all, in order to express the temperature of a substance in terms of the energy per gram molecule we have to select a different mass for each substance—for example, 2 gm. of hydrogen, 20 gm. of methane. To say that this results in $\left[\frac{\text{energy}}{\text{mass}}\right]$ for the dimensions of temper-

---

* A more general statement is also given to cover any substance, as well as radiation (Duncanson, 1941).

ature is equivalent to deducing the dimensions of intensity of magnetization to be $\left[\frac{\text{magnetic moment}}{\text{mass}}\right]$ because for iron it is the magnetic moment of 7·8 gm. or for mercury the magnetic moment of 13·6 gm. It is obvious that these masses of iron and mercury have been chosen in order to obtain unit volume (1 c.c.) of each substance, so that this is merely another way of defining intensity of magnetization as magnetic moment per unit volume. Similarly, the masses of hydrogen and methane have been chosen to obtain the same definite number of molecules in each case, so that temperature defined in terms of energy per gram molecule has the dimensions of [energy]. The fact that the particular number of molecules is not known exactly is immaterial as long as there is a means of selecting the same number for each different substance.

Returning to Benham's deduction that temperature has the dimensions $[L^{-1}]$, we see that this will not hold if our previous contention is admitted that it is an unnecessary complication to introduce the constant N into the force equation. Moreover, not only is it a complication but, as we have already seen, it introduces two unknowns, namely F and N, into the one equation. At first sight it appears that Benham has devised an independent method for determining the dimensions of N. However, on examination it is found that, even after making several assumptions, he adopts a procedure which in the end amounts to making an arbitrary choice for the dimensions of N. For example, one of the alternative possibilities for its dimensions is rejected because the dimensions of temperature deduced therefrom, namely $[ML^{-1}]$, appear to have " no very obvious physical significance ". Such a procedure does not seem to be sufficient to justify the introduction of N.

This matter has been dealt with in some detail as it demonstrates the importance of adequate definition of physical quantities, and the unnecessary confusion of dimensions that results unless this precaution is carefully observed.

§ 4. INTENSIVE AND EXTENSIVE MAGNITUDES

Tolman (1917) has pointed out that temperature is a quantity of intensive magnitude, while energy is of extensive magnitude. This may be raised as an objection to our definition of temperature in terms of the mean energy of a molecule. Let us recall that a quantity of extensive magnitude is such that if it has a value a_1 for body 1, and a_2 for body 2, then it will have a value $a_1 + a_2$ when the bodies are taken together. Intensive quantities are not additive in this manner.

It is interesting to examine a little more closely the nature of some of these intensive quantities. It will be found that intensive quantities include some which are defined in terms of certain amounts * (it may be a certain number, length or mass, or, what comes to the same thing, some kind of average) of the matter to which the quantity refers, and it is easy to see then how these quantities cannot be extensive, as their very definition precludes this possibility. For example, the density of a substance is defined as the mass of a certain volume of that substance, and even if two bodies of this substance are added together,

---

* " Amount " is used rather than " quantity " to avoid using the latter with two distinct meanings.

their density remains the same, as it is still the mass of the same sample volume. It is easy to see that the same applies to such quantities as thermal capacity per unit mass and intensity of magnetization (magnetic moment per unit volume). Temperature is such a quantity. It is defined in terms of the extensive quantity, energy, but is prevented from being extensive itself by the fact that the energy is to be that of a stated number of molecules, or their mean energy, according to the definition preferred. So there is no inconsistency in temperature, which is an intensive quantity, having the same dimensions as that of the extensive quantity, energy.

NOTE ADDED IN PROOF. Since this paper was communicated, the author has had the opportunity of seeing two papers which cover, to some degree, similar ground to the above note, one by Dingle (*Phil. Mag.* **33**, 321, 1942) and the other by Guggenheim (*Phil. Mag.*, in the press).

REFERENCES

BENHAM, W. E., 1942. *Proc. Phys. Soc.* **54**, 121.
BROWN, G. B., 1941. *Proc. Phys. Soc.* **53**, 418.
DUNCANSON, W. E., 1941. *Proc. Phys. Soc.* **53**, 432.
GUGGENHEIM, E. A., 1942. *Proc. Phys. Soc.* **54**, 290.
HANSEL, C. W., 1942. *Proc. Phys. Soc.* **54**, 151.
TOLMAN, R. C., 1917. *Phys. Rev.* **9**, 237.
YARNOLD, G. D., 1942. *Proc. Phys. Soc.* **54**, 46.

DISCUSSION

On papers by C. W. HANSEL (*Proc. Phys. Soc.* **54**, 151, 159, 164 (1942)).

Sir CYRIL ASHFORD. I am on Mr. Hansel's side in preferring $P=ma$ to $P/W=a/g$ as a teaching method—though that question seems to me of less importance than the fundamental one of whether school kinetics should be treated as a deductive or an inductive science. My experience goes to show that it can be plausibly presented in the latter way even to post-certificate schoolboys, but I rather think that it is only the pick of them that really appreciate the logical difficulties.

Mr. J. L. BRERETON. My position can be briefly summarized as follows :—

(1) I am not in agreement with the report of the Mathematical Association on the Teaching of Mechanics when it states that definition of mass (through inertia) should be avoided in the early stages by use of the formula $P/W=a/g$, because, from discussions with university teachers, I believe many pupils fail to change over effectively later to the more significant formula $P=ma$.

(2) Further, we have in secondary schools not only to train children to become future teachers or research workers, but also to see that those concepts which have helped to build our industrial world—mass, force, velocity, energy, temperature, etc.—become part of the ordinary awareness of the people. An example is the crying need, in relation to flying, for a thorough understanding of relative velocity, a conception which we have notoriously failed to inculcate. One of the most important of these concepts, at any rate for the chemist, and in many commercial transactions, is " quantity of matter " expressed in units of mass and measured, as Mr. Hansel puts it, in " any convenient manner ". I cannot see that either mass or relative velocity is more difficult than the relation between heat and temperature, which most children tackle early in their school course and have little difficulty in grasping.

(3) I do not consider that a case has been made against the definition of mass as quantity of matter. For bodies of the same material, quantities of matter can be compared by means of volume or by chemical equivalence (as in volumetric analysis). For bodies of different material, quantities of matter can be compared by means of (i) weight, (ii) inertia. (*See* Clerk Maxwell, *Matter* and *Motion,* Art. 46.)

(4) This idea of using one quantity as a measure of another is important in itself. For example, in a car running at constant speed, distance travelled can be measured on a watch ; children learn to compare areas of cardboard by weighing, pressures by heights of a column of mercury, temperature by volume of gas, force by extension or twisting of springs, etc.

(5) My only quarrel with Mr. Hansel is in his generalization that a beam balance measures mass and a spring balance measures weight. Both instruments have to be used to measure both quantities, and the student must understand this. I cannot agree with his " mass in water " in an experiment on the principle of Archimedes. Surely the child can understand that *in certain circumstances* mass is proportional to weight and weight to mass.

(6) The objection is often raised that it is difficult to explain in simple language what one means by quantity of matter. But this applies equally to most physical quantities. Length, time, force can only be explained and defined in terms of the method of comparing two lengths, or two times, or two forces, as the case may be.

(7) Then there is the difficulty that the quantity of matter in (say) an electron varies with its speed ; but so, after all, does the length of a rod and an interval measured on a clock. It must be admitted that at present we are not in a position to introduce the ideas of the special theory of relativity in elementary mechanics classes.

(8) The fact that inertial mass and gravitational mass (*a*) are both additive, (*b*) are identical to the limits of our measurement, is strong justification for saying that mass is more fundamental than either of these manifestations, and that Newton's identification of it with quantity of matter was correct. After all there is no other property of bodies which is additive in the way in which mass is.

(9) I suspect that much of this attempt to soft-pedal the association of mass with quantity of matter is traceable to Ernst Mach, who was one of the first to give the definition of mass which assumes that the weight of a body is proportional to its inertia, but who had to discredit the concept quantity of matter in order to support his (to my mind untenable) theory that matter was nothing but a sensation in his brain. It is significant that in his scientific writings he was forced repeatedly to throw this idealist theory overboard.

Mr. K. J. CARPENTER. In his paper on the definition of mass, Mr. Hansel glosses over the fact that gravitational force on a body, in a field of given strength, is in a constant proportion to the mass, where mass is found by comparison of inertia with a standard body (this definition of mass being independent even of the existence of such a thing as gravity).

This is a very remarkable phenomenon, and no more self-evident than to suppose the electrostatic and magnetic pole strengths of a body to be determined by its mass.

I therefore submit that the beam balance does *not* measure mass, but compares weights, because the opposing couples about the fulcrum are due to gravitational forces, which are present owing to the *weights* of the bodies.

In an article on physics fundamentals, could not some experimental evidence of the proportionality of mass and weight be included ?

The period of a compound pendulum may be written $\sqrt{\frac{Mk^2}{Wl}}$ (where M is its mass and W its weight). If, at a given place, the M/W ratio is always found constant, the proportionality will have been *approximately* checked.

Prof. J. A. CROWTHER. Though it is hardly possible to add materially to the very full and, I may add, extremely able discussion of the teaching of mass and weight contained in the paper by Mr. C. W. Hansel, I should like to take the opportunity of expressing my agreement with his main conclusions. The chaotic state into which the teaching of mechanics has been thrown by well-meant, but misguided, attempts to placate the "practical man" following on the B.A. Report of 1907 is well known to all examiners of School Certificate physics. The mental confusion of these unfortunate candidates is as pitiable as it is unnecessary.

The trouble arises from two main causes. In the first place there has been a tendency to regard mechanics as a subject in itself, divorced from the great body of physical science of which it forms a small, though fundamental, part. Any system which necessitates such statements as "the mechanical equivalent of heat is 772 ft.lbs. per B.T.U. at Manchester" is automatically ruled out of court as a suitable medium for scientific thinking. Similar considerations apply to the suggestion, which Mr. Hansel appears to accept, that the gravitational system of units should be rationalized by the introduction of a formal unit of mass, the "Slug". A suggestion which would appear to demand such statements as "unit mass (1 slug) of hydrogen contains $2 \cdot 15 \times 10^{22}$ molecules at Wigan, and a sixth of that number on the surface of the moon", is not calculated to appeal to any serious scientific thinker.

That such a system should have found its way into our schools is due, I believe, partly to the confusion existing in so many quarters between education and technical training, and partly to the belief, often strongly held, though not invariably expressed, that if the two are distinct, the latter is fundamentally important ; the former being merely a luxury, like whisky or cigars, not positively harmful unless indulged in to excess. Hence the stress on copying the methods of the "practical man". Education should precede, not accompany, technical training ; and a properly educated student of mechanics will have no difficulty in realizing (will, in fact, have already realized) that where high accuracy is not important, the weight of a pound mass is a handy practical equivalent to a force of 32 poundals. If he finds that the eviscerated equation $P/W=a/g$ simplifies his calculations (which is doubtful) he will at least use it with understanding. I believe that gravitational units should be eliminated from our elementary teaching, or at any rate relegated to an appendix at the end of the text-book. No other single reform would do more to increase the efficiency of scientific education in our schools and universities.

Mr. A. F. DUFTON. In his paper on Mass, Mr. Hansel is unrestrained in his criticisms of those who in the past have given consideration to the teaching of elementary mechanics : I note in particular his remarks about the report of a most competent and distinguished committee of the British Association. Clause 19 of this report he stigmatizes as an ill-conceived suggestion, and, apropos of Clause 20, he asks "Can anything be more fatuous ?"

This question deserves an answer, and it will be found, I think, in § 9 of Mr. Hansel's paper, where, in an example of the Principle of Archimedes, the mass of a solid in air is 100 grams and the mass of the same solid in water is blandly stated to be 80 grams. Is not this the quintessence of fatuity ? Where are the other 20 grams ?

Would not all the pother about gravitational units disappear if the pound weight were used in conjunction with the pound mass and g were recognized as the corresponding unit of acceleration ?

1 *lb.* weight produces 1 g acceleration, in mass 1 *lb.*

The gravitational system is a local system, and it is not the mass and the density of a body which vary from place to place but g, the unit of acceleration.

Mr. D. G. A. DYSON. I do not think that mass is a very easy idea to grasp really accurately, but neither is any fundamental concept ; I certainly should not care to *explain* the quantities quoted as more difficult—length and time. Was Mr. Lauwerys very helpful in quoting from a number of distinguished men of various nationalities to show how difficult they found the clear understanding of mass ? In any case, can the schoolboy be expected to have quite the clarity and accuracy demanded by scientific giants in their treatment of such fundamentals ? I think the idea of inertia gives the meaning of mass quite as clearly as is necessary, especially after reference to trolley experiments and everyday experience. In any case, the idea of mass must be used sooner or later, and it is no solution to say that no one really understands it.

As far as $P/W=a/g$ is concerned, I do not understand why it was ever introduced. In the first place, I do not agree with Mr. Ghey's implication that a boy finds ratios easy to understand ; my experience has been that this is one of the most difficult of elementary ideas for him to grasp. In the second place, the experimental demonstration of Newton's second law and its application to the definition of force units is perfectly easy to understand. I suppose ratios come into this derivation too, but not so fundamentally as in the P/W formula. In the third place, this P/W formula appears to tie everything to the idea of weight, whereas the $F=ma$ of Newton's second law is of universal application, and, once grasped, absolutely foolproof. I have felt for a long time that many of the new teaching methods designed to make the work easier for the pupil only do so (if at all) by evading the really important points.

I am not sure that I would go the whole way with Mr. Hansel in not using the gram weight, etc., in the early stages of mechanics, before the introduction of the laws of motion. I do not feel that their subsequent dismissal is very serious. On the other hand, I am bound to admit that his view-point is really sounder, and I am quite open to conviction. It is just whether one can be really strictly logical with children, whose minds are not able to grasp the whole range of the subject at once.

I was very sorry that the author did not have more support in the discussion, but no doubt there must have been many who did not speak, simply because they were in agreement with him.

Mr. G. W. E. GHEY. " A resultant force of 10 tons weight acts on a mass of 20 tons. What is the acceleration " ? Surely the simplest solution is the application of $P/W=a/g$, as follows : Since 20 tons weight acting on 20 tons gives acceleration g, therefore 10 tons weight acting on 20 tons gives acceleration $\frac{1}{2}g$.

Mr. G. F. HERRENDEN HARKER. The most illuminating physical approach to Ohm's law is from energy considerations.

Consider a portion AB of an electric circuit consisting of a stationary metallic conductor of resistance r_{AB}. Then if V_A denotes the potential of A, V_B that of B, and if this potential difference V_A-V_B volts maintains a current of i amp. between A and B during an interval of t sec., thus transferring a quantity q $(=it)$ coulombs of electricity from A to B, the electrical energy expended between A and B will be given, in joules, by $w_{AB}=(V_A-V_B)q$.

If, further, the only physical phenomenon occurring between the points A and B is an evolution of heat within the conductor, the heat energy liberated will, by Joule's law, be given by $w_{AB}=r_{AB}i^2t$, in which, if w_{AB}, i and t are measured in the units stated above, r_{AB} will be given in ohms.

Equating these two expressions for w_{AB} gives

$$V_A-V_B=r_{AB}i.$$

This is Ohm's law, which is accorded such exaggerated prominence in elementary texts. *It is only valid for a portion of an electric circuit in which the sole physical happening consists in an integral conversion of electrical energy into heat energy.* It thus provides the key to the solution of problems concerned with electric lamps, radiators, irons and toasters, which, however, by no means cover the entire or even the most important among the practical applications of electricity.

The portion of the circuit between A and B may include an apparatus capable of transforming some part of the electrical energy supplied into something other than heat. Thus, if the apparatus is an electric motor, some of the electrical energy supplied to it will appear in the form of mechanical work, If, of the total quantity W_{AB} of electrical energy supplied during a time t between the terminals A and B, a portion, w_{AB}, is transformed into heat, while the remainder, w'_{AB}, is converted into other forms, so that $W_{AB}=w_{AB}+w'_{AB}$, then, as before,

$$(V'_A-V'_B)i't=r'_{AB}i'^2t+w'_{AB},$$

i.e.,
$$V'_A-V'_B=r'_{AB}i'+w'_{AB}/q'.$$

The simple version of Ohm's law is thus no longer valid, and the second term on the right-hand side required to complete and generalize it, representing, as it does, the energy converted by the apparatus into forms other than heat per coulomb of electrical energy supplied, is normally of much greater importance than the first term on the right-hand side, representing the internal ohmic voltage drop within the apparatus. This second term on the right-hand side, obtained by dividing a number of joules by a number of coulombs, will be expressed in volts. The corresponding voltage $v'=w'_{AB}/q'$ can be regarded as acting in opposition to the applied voltage, so that, in formal analogy with the previous relation, we can write

$$(V'_{\mathrm{A}}-V'_{\mathrm{B}})-v'=r'_{\mathrm{AB}}i'.$$

Note that if the resistance between A and B remains unchanged, i.e. if $r'_{\mathrm{AB}}=r_{\mathrm{AB}}$ (as, for example, if the " conductor " considered in the first case is an electric motor with its shaft clamped, and in the second case is the same motor with its shaft free) and if the P.D. applied between the terminals A and B is the same in both cases, the current i' in the second case will be less than the current i in the first case. Alternatively, if the current is the same in both cases, the P.D. $V'_{\mathrm{A}}-V'_{\mathrm{B}}$ in the second case will be greater than the P.D. $V_{\mathrm{A}}-V_{\mathrm{B}}$ in the first case.

When a generator of e.m.f. E volts maintains a current of I amp. in a circuit during an interval of t sec., the total electrical energy supplied by the generator will be given, in joules, by

$$W=EIt,$$

and will reappear in various other manifestations of energy throughout the entire circuit. Thus, if R denotes the resistance of the generator, r that of the remainder of the circuit, and if the whole of the electrical energy supplied by the generator suffers conversion into heat, then

$$EIt=(R+r)I^2t,$$

i.e.,

$$E=(R+r)I.$$

Note that as $v=rI=E-RI$, the P.D. between the terminals of the generator when furnishing current is less than its e.m.f. by the internal ohmic voltage drop within the generator itself.

If part of the electrical energy made available by the generator appears in some form of energy other than heat, and if, during the interval t, its amount is denoted by w', then

$$EI't=(R+r)I'^2t+w',$$

i.e.,

$$E=(R+r)I'+w'/I't.$$

The second term on the right-hand side having the physical dimensions of a voltage can be denoted by V', so that we can write, in formal analogy with the previous relation,

$$E-V'=(R+r)I'.$$

With regard to the enunciation of Archimedes' principle, I should like to urge the complete suppression of any mention of " apparent loss of weight " and the substitution of " upthrust " in its place.

Mr. J. A. LAUWERYS. Mr. Hansel expresses the views of the school of thought described in the Mathematical Association *Report* as " being dominant at the end of the last century but now gone out of fashion ". Since I take it that Mr. Hansel wishes to convince us of the soundness of his views, he did well to start as he did. For in the three points which he asks us to assume, he has covered up in ambiguity and obscurity all the difficulties of the topic in such a way that, having once gone wrong there, one can never get straight again. In consequence one may be misled into accepting his conclusions.

(1) Take his first point. " It is advisable to teach a student the conception of mass and not to ignore it completely." What is meant by " teach the conception of mass " ? Give him some sort of an idea ? Or teach him enough to enable him to solve difficult problems—say about water-carts with holes in their bottoms running down a hill ? Are " teach " and " ignore " the only alternatives ? Lastly—most important—does the author here mean " inertia " when he says " mass " ?

(2) If we grant that mass, defined as quantity of matter, is a measure of inertia, the question might arise as to when the word is first to be introduced. But if we use $P/f=W/g$, we do not have to change over to any new system. At the appropriate moment, when the pupil has reached the appropriate stage of intellectual development—say at Higher School Certificate stage—we extend. As Prof. Alfred Lodge wrote in the *Math. Gazette* (Feb. 1938): " For beginners I am convinced that Newton's second law is best given as $P/W=f/g$, i.e. accelerations are proportional to the forces acting on a given body, actually or potentially. The equality $P/W=f/g=$ a constant for a given body is more abstruse and should be delayed a bit. $W/g=m$ is chiefly a matter of *quality*. W and g are both vertical vectors ; m is a scalar quite directionless. Important, but too advanced for youngsters : that is the reason for delaying a dissertation on M."

Incidentally, in the same number of the *Math. Gazette*, is printed a presidential address by the late Prof. L. N. G. Filon on this very topic; it concludes as follows :—" It seems at any rate remarkable that after all the fuss we have made about the fundamental distinction between mass and weight and the criticisms levelled at the engineer for his systematic use of W/g, it should turn out that the engineer was right after all, and that incidentally Newton's amazing intuition had, in that disregarded definition, hit upon the inevitable truth."

(3) What are these "unnecessary, unwanted, irrational units" ? The weight gram or the pound weight ? ; the cm. ? ; the foot ? The whole point of $P/W=f/g$ is that it does not bring in new units.

(4) Mr. Hansel's remark about the confusion in the minds of students in this country and in the U.S.A. has point only if the confusion does not exist in other countries. Let us then turn to Karl Hahn's monumental *Methodik des Physikalischen Unterrichts*. He spends page after page (4+18 quarto) deploring how hard it is to make pupils get their sums right, how muddled they are, how difficult the whole thing is, and his solution is : in the lower forms, by all means, try to make clear the idea of *Inertia* and distinguish it from *Weight*. After the age of 16, approach gradually the concept of mass—experimentally, e.g., via Attwood's machine, etc.

Henri Poincaré also spends a lot of time on this question. *Inter alia*, he asks how mass is to be defined, discusses it at length and concludes : " Mass is a numerical coefficient which mathematicians find it convenient to use in their equations." He deplores the bad teaching of mechanics in France. Pupils, he says, think of forces as arrows with which one draws parallelograms. How he wishes that it was taught as well as in England ! There they teach it sensibly—as an experimental science : " Naturellement, les Anglais, comme toujours, ont raison."

(5) " A beginner can understand mass with less difficulty than he can understand length and time."

" Teach the distinction between mass and weight early in the science course."

These questions have been discussed at great length by many eminent mathematicians and logicians, e.g. by Ernst Mach in *The Science of Mechanics*, Karl Pearson in *The Grammar of Science*, Henri Poincaré in *Science et Methode* and *La Science et l'Hypothèse*. (See also Prof. Watson's *On understanding Physics* and Prof. Filon's address in *Math. Gazette*, Feb. 1938.)

I wish I could describe in detail what their conclusions are, but they all agree

(*a*) how very difficult the topic is, and how muddling to beginners ;

(*b*) how Newton's genius shines out in his treatment of the question ;

(*c*) how nearly impossible it is to give a satisfactory definition or explanation of what is meant by mass ;

(*d*) that the only satisfactory approach is via Newton's laws of motion.

Thus, Karl Pearson says : " I must now ask the reader to follow me in a rather more difficult investigation . . . . $m_1f_1=m_2f_2$, etc. Mass accordingly as the ratio of the numbers of units in two accelerations is a conception which can easily be appreciated. It is in this manner that mass is invariably determined scientifically. Yet the reader will frequently find mass defined as ' the quantity of matter in a body '. After our discussion the reader will easily appreciate how idle is a definition of mass in terms of matter ".

Again, from E. Mach :—" The expression ' quantity of matter ' is not adapted to explain and elucidate the concept of mass, since that expression itself is not possessed of the requisite clearness. It did not occur to Galileo that mass and weight were different things Huyghens, too, in all his considerations put weights for masses. . . . Physicists were not led to form the concept *mass* till they made the discovery that the *same* body can by the action of gravity receive different accelerations."

Then we have L. N. G. Filon :—" Newton, after talking round the question, defines mass by weight. Curiously enough, this reference to weight in the *Principia* seems to have been almost completely ignored, and for over two centuries we find the purely verbal definition of mass by ' quantity of matter ' repeated without explanation. If now we proceed in logical manner, and leave weight out of consideration for the present, ' mass ' in the second law is merely an undetermined constant belonging to a body, and it appears that all that the second law can tell us is how to compare forces."

Finally, H. Poincaré : —" The definitions proposed are sheer metaphysics."

(6) It is to be noted that the force equation $P/W=f/g$ has been proposed precisely by those who have thought carefully about the difficulties in the concepts and have been struck by the difficulties which the pupils meet when attempts are made to teach the distinction between mass and weight early in the course.

What they have wished to avoid is what the adoption of Mr. Hansel's ideas would ensure, viz., giving pupils the illusion of knowledge or substituting confusion for ignorance.

The best course to adopt is to follow the advice given in the Mathematical Association *Report*, thus making co-operation easier between teachers of science and of mathematics.

This means : (*a*) being careful to use words properly ourselves, as Mr. Hansel says, and (*b*) making clear at some stage the distinction* between inertia and weight, but not defining mass ; (*c*) leaving mass and absolute units until after school certificate.

(7) (*a*) For the rest, I would be quite willing to adopt the word *slug*, though as Mr. C. O. Tuckey says plaintively in a letter published in the *Math. Gazette* (1938, p. 282), " the name ' slug ' sounds idiotic and unscientific ". (Why do text-book writers hold the slug in contempt ? I believe the answer to be simple, though silly.)

(*b*) The $P/W=a/g$ formula has been preferred not only by B.A. Committees, but by most committees.

(*c*) Lastly : Do not imagine that ordinary weighing machines compare *masses*. They do not. They compare *weights*. Consider what would happen if a balance was placed in a highly heterogeneous gravitational field. We can use them to compare masses only after we have experimentally demonstrated Newton's law that $m \propto W$ (the foundations of Einstein's later work on General Relativity).

There is always a danger of treating dynamics chiefly as a mathematical discipline enabling the learners to solve formal problems and exercises ; as A. N. Whitehead says : " My own criticism of our traditional educational methods is that they are for too occupied with intellectual analysis, and with the acquirement of formularized information. We are too exclusively bookish in our scholastic routine. The general training should aim at eliciting our concrete apprehensions and should satisfy the itch of youth to be doing something. In the Garden of Eden, Adam saw the animals before he named them : in the traditional system, children named the animals before they saw them."

My own view of the teaching of dynamics is that it is a study of the real world, an attempt to explain it by generalizations which increase in scope as the experience of the learners grows, and where the rationalizations must increase in abstractness only as their intellectual grasp deepens. Hasten the process unduly and you help to build up an academic, scholastic mind, not a scientific one.

Mr. Rottenburg. To be inconsistent in teaching is always a great mistake, and we have been inconsistent when we have used the word *mass* for a property of a body in discussing linear motion, and the word *inertia* for the same property when discussing rotational motion. Newton was quite clear in teaching that *mass* was compounded of *volume* and *density*, i.e. that it was quantity of matter. He also describes the force a body exerts when it has its motion altered as the *vis inertiae*—in other words, that the inherent property that sets up the force is *inertia*. But even without Newton the use of two different words in the two cases remains a mistake. Mr. Hansel says we must teach " mass ", clearly—but what happens then ? The pupil goes out after the lesson and meets the word right and left—(the masses ; the mass of the electorate ; masses of this or that ; mass production ; mass psychology)—but used in a sense that is never what he has been taught. And the reiterated meaning (namely " a quantity of ") will surely win. Every pupil, too, is liable to be struck by the fact that, even in his teaching, the inconsistency of using *inertia* in one set of cases and *mass* in the other is wrong.

---

* How has the confusion arisen in this paper ? The key is found in Mr. Hansel's sentence, " Mass is a property of things ever present in daily experience "; he means inertia, of course.

When I first started to use the word *inertia* in teaching I found many pupils relieved when I informed them that it was the same property as had been taught them as *mass*. I also found that the word *inertia* helped me to keep my own ideas clear. One can then go on and explain that either *weight* or *inertia* can be used to measure Newtonian mass. If a man sits on the edge of a merry-go-round at rest with a balance and " weights " or bits of metal marked 1 lb., he can measure a mass or quantity of matter by balancing the bits of metal against the " mass " in the other pan. If he now starts up the merry-go-round, and it reaches a speed where the scale-pans fly out almost horizontally, he can still compare and balance the bits of metal and the " mass " but he will now be using their inertia to compare the masses, whereas, while at rest, he compared them by their weights.

Prof. G. STEAD. As regards paper I, on Ohm's law, I am in general agreement with the conclusions set out on page 159, except that it seems to me to be too late in the day to make the changes in the meaning of *ampere*, *volt* and *ohm* suggested in § 5. Such changes would result in hopeless confusion and would do more harm than good.

I know from experience that the experimental illustration of Ohm's law described on pages 156–7 does, in fact, help beginners to grasp the essential point of the law, but I think that teachers should be perfectly honest with their pupils and tell them plainly that the experiment is in no sense a *proof* of Ohm's law. The measurements of P.D. are made with a moving-coil voltmeter, which itself depends on the truth of Ohm's law, and, therefore, all the experiment really proves is that the instrument maker has calibrated his meter with reasonable accuracy. The experiment should really be done with an electrostatic voltmeter, but this, of course, is impracticable, at least for beginners.

The extracts, in the paper on specific heat, from the writings of various physicists, seem to indicate that the whole question of the definitions of thermal quantities needs clarifying. It seems to me that it is not so much a matter of choosing between two alternative definitions of specific heat as of recognizing that nearly all of us are guilty of using the single term *specific heat* to denote two essentially different things. It is usually said that the water equivalent of a calorimeter of mass m grams and specific heat s is (approximately) ms gm., and that the thermal capacity of a body of mass m gm. and specific heat s is ms calories per degree. Clearly the first s is a pure number and the second is a dimensional quantity measured in calories per gram per °c. Hansel is, of course, correct in saying that, in the equation $H=ms(t_2-t_1)$, s cannot be a pure number, but he seems to have overlooked the fact that it is not this s which is measured either by the method of mixtures or by the method of cooling. If s is a dimensional quantity, the usual equation for the method of mixtures, viz.,

$$ms(t_2-t_1)=m_0(t_1-t_0),$$

is clearly wrong dimensionally, the correct equation being

$$ms(t_2-t_1)=m_0s_0(t_1-t_0),$$

where s and s_0 are the thermal capacities per unit mass of the solid and the water respectively, and from this we find

$$s/s_0=m_0(t_1-t_0)/m(t_2-t_1).$$

Similarly, with the method of cooling, we compare a liquid with water and obtain

$$ms(d\theta/dt)=m_0s_0(d\theta/dt)_0,$$

which again gives us s/s_0 and not s. Thus, the elementary methods of determining specific heat give the result as a ratio. If anyone doubts this, I would point out that in mixture and cooling experiments both masses and temperatures can be expressed in any units whatever without affecting the result.

Electrical methods give, in the first place, a measurement of the thermal capacity per unit mass expressed in ergs (or joules) per unit mass per degree, and a conversion of the result to calories per gram per °c. involves a determination of J. If we also do this electrically, we have in effect determined the thermal capacity per unit mass (*a*) of the given substance, (*b*) of water, and might well write

$$Jms\,(\theta'-\theta) = Eit,$$

$$Jm_0s_0(\theta_0'-\theta_0) = E_0i_0t_0,$$

so that

$$s/s_0 = Eit(\theta_0'-\theta_0)m_0/E_0i_0t_0(\theta'-\theta)m.$$

Thus we obtain s/s_0 as a pure number, and do not require to know the value of J.

It seems to me to be a matter of little importance whether the term *specific heat* is used to express a ratio or a dimensional quantity, but it is unfortunate that it should be used indiscriminately in both senses. Nor can the trouble be overcome by considering only one aspect of the matter and deliberately shutting our eyes to the other. The truth is that two different names are needed, as in the strictly parallel case of specific gravity and density. Until these have been agreed upon I think there is a case for defining specific heat as a ratio, and using the term *thermal capacity per unit mass* for the dimensional quantity. Specific heat would then correspond to specific gravity, and would be the constant actually measured in " specific heat " determinations. In any case, I think that all definitions of the following type should be absolutely rejected :—

" The specific heat of a substance is the amount of heat required to raise the temperature of 1 gm. of it through 1° c."

This associates the entirely nebulous term " amount of heat " with the exact terms 1 gm. and 1° c. One might just as well define the density of a substance as the amount of matter contained in 1 c.c. of the substance.

Incidentally, I feel that there is a good deal of confused thinking even on the subject of specific gravity and density. Some teachers have suggested that the term specific gravity should be abolished, as unnecessary. The fact is that nearly all the common methods of experiment (e.g. specific-gravity bottle, hydrostatic balance, Hare's apparatus) determine specific gravities and not densities. Some teachers even go so far as to call a specific-gravity bottle a *density bottle*, though it is obvious that only specific gravities can be found by this method. Anyone who doubts this should provide himself with a balance, a specific-gravity bottle and a set of English weights and try to find the density of a liquid in pounds per cubic foot (without assuming someone else's result for the density of water in these units). The accurate determination of the density of a liquid is very much more difficult than a measurement of its specific gravity.

To return to the question of specific heat, I woud suggest the following as a consistent and logical approach to calorimetry :—

(*a*) Choose an arbitrary unit of heat, and any convenient scale of temperature.

(*b*) Define the *thermal capacity* of a body as the quantity of heat required to raise the temperature of the body one degree. Thus, thermal capacity is expressed in heat units per degree.

(*c*) Denote the *thermal capacity per unit mass* of any substance by c, and the thermal capacity per unit mass of some standard substance by c_0. Then c and c_0 are measured in heat units per unit mass per degree.

(*d*) Define the *specific heat* s of a substance as

thermal capacity per unit mass of substance/thermal capacity of unit mass of standard.

Thus $s=c/c_0$, and is a pure number.

On this basis, the thermal capacity of a body of mass m and specific heat s is msc_0 *heat units per degree*, and the " standard "-equivalent of a calorimeter of mass m' and specific heat s' is *$m's'$ units of mass.* In a specific-heat determination by mixing a hot solid with the cold standard substance we have an equation of the type

$$msc_0(t_2-t_1)=m_0c_0(t_1-t_0)+m'sc_0(t_1-t_0)\ ;$$

whence $$s=(m_0+m's')(t_1-t_0)/m(t_2-t_1).$$

Thus we obtain s as a pure number without requiring to know c_0, and the masses and temperatures may be expressed in any units whatever without affecting the result. The method of cooling and electrical methods can be treated in the same way.

It may be, and generally is, convenient to choose our units so that $c_0=1$, but this is not essential any more than we are obliged to use units in which the density of water is unity. It may be noted that if we use absolute c.g.s. units, the value of c_0 for water is not unity, but $4{\cdot}2\times10^7$ ergs per gm. per °c.

I agree with Mr. Hansel, in his paper on Mass, that students of average intelligence are quite capable of understanding absolute units, and of learning the correct use of the terms *mass* and *weight*, and I support his view that the second law of motion should be applied in the form $F=ma$. The equation $F/W=a/g$ is neither necessary nor desirable, and momentum, kinetic energy, and moment of inertia should be expressed as mv, $\frac{1}{2}mv^2$, mk^2, and not as wv/g, $wv^2/2g$, wk^2/g. It seems to me to be fundamentally wrong to introduce some local value of g into formulae and expressions which have nothing whatever to do with gravitation, and which are presumed to be of general applicability, not only on the earth, but on any other celestial body, or, indeed, at any point in space. In a consistent system of engineering units a special unit of mass is undoubtedly required, and although the name *slug* seems to me to be peculiarly ugly, I agree that it is much better to use it than to confuse our students by pretending that the engineer's unit of mass is the pound, and then trying to put matters right by introducing chaotic expressions and equations. But is it really necessary for students of physics to be taught engineering units as a separate system ? Is it not sufficient to give a general instruction to beginners to convert everything into absolute units before substituting in a dynamical formula or expression ? Students soon acquire the habit of converting miles per hour into feet per second before substituting in a kinematical equation, and it is equally easy for them to learn to convert pounds weight into poundals before substituting in $F=ma$.

I think that Hansel's statement that a beam balance measures *mass* and a spring balance *weight* is too sweeping and needs further discussion. The fundamental formula for a beam balance in equilibrium is $m_1g_1d_1=m_2g_2d_2$, which expresses the fact that the weights on the two sides have equal moments about the support. The balance is constructed so that $d_1=d_2$, and experiments with pendulums show that g is the same for all kinds of matter at a given place, so that ultimately $m_1=m_2$, and the balance does in this way determine equality of mass. In a spring balance the weight mg is balanced against the tension in the spring, but the graduations are obtained by hanging standard " weights " on the spring, so that we are again comparing m_1g_1 with m_2g_2, but this time with the spring as intermediary. Thus, if $g_1=g_2$, we again have $m_1=m_2$, and the graduations represent masses, not weights. Thus it seems to me that there is no real distinction between " weighing " by beam balance and by spring balance. The spring balance is, of course, only calibrated for one value of g, and, therefore, gives different readings in different places, but I see no justification for using " weight " to express a spring-balance measurement and " mass " for a beam-balance measurement. I think that, in general, mass should be used in both cases.

I must also criticize Hansel's example on the hydrostatic balance on page 172. He writes : " Mass of solid in air = 100 gm." " Mass of solid in water = 80 gm." Surely this is completely false, or at least entirely misleading ? How can the *mass* of a solid be any less in water than in air ? It is completely contrary to Hansel's own statement earlier in the paper that teachers should stress the constancy of the mass and the variability of the weight of a given body. Would Hansel say that the mass of a body floating in a liquid was zero ? Surely the hydrostatic balance is measuring the tension in the supporting thread, not the mass of the suspended body ? I suggest that a consistent statement of Hansel's problem would be :—

Tension in supporting thread when body is in air = 100 gm. wt.
„ „ „ „ „ water = 80 gm. wt.

I think, however, that we might be permitted to express the same thing more shortly, if slightly less accurately, by saying, "Weight of body in air = 100 gm."; "Weight of body in water = 80 gm."

Finally, I should like to express my warm thanks to Mr. Hansel for the valuable work he has done in collecting together the large quantity of material contained in these papers, and in setting it out in a form which cannot fail to be of great assistance to all teachers of elementary physics.

Dr. G. T. P. **Tarrant.** For the past three years I have been teaching my boys at their earliest stage to insert in their formula not only the numbers of the physical quantities concerned but also the units as well. Thus, Density is $\frac{24 \text{ gm.}}{3 \text{ c.c.}} = 8$ gm. per c.c. ; $I = \frac{\pi \rho r^4 l}{4} = \frac{\pi\, 2{\cdot}4 \text{ gm./cm}^3\, (1{\cdot}0 \text{ cm.})^4 \times 10 \text{ cm.}}{4} = \pi \times 6 \text{ gm.cm}^2$ The boys then solve the unit as well as the numerical part. The advantage of this system is that it draws attention to the units, shows them immediately if they are making any mistake by confusing, say, inches and cm., makes them clearer on the difference between density and specific gravity and enables them to present their result with the correct unit attached.

It has also just struck me, though I have not tested it, that the difficulty of forces in lb. wt. or in poundals would largely disappear if this method were adopted and if, instead of lb. wt., we simply said lb.-g. and cancelled out the lb. with lb. but left the g standing.

The only feature which introduces any difficulty is the names that are given to composite units—it is generally better for the boys to write lb.cm.sec.$^{-2}$ than poundals.

Let us now turn from this to the question of the dimensions to be assigned to electrical and thermal units. We are interested in dimensions because they are of use to us. With their help we can check equations, solve problems and convert values in one set of units to values in another set. In mechanics we start with length and time and find the units of acceleration. We then meet the equation $P = mf$, one equation containing two new quantities P and m. We must fix one arbitrarily—mass. It would be silly to say that we won't worry about the dimensions of this new quantity, mass. If we did we would lose a portion of the richness coming from the theory of dimensions. We could solve problems containing only two unknowns instead of the normal three. We could convert quantities from one set of units to another only if there was a change in the units of length and time, and could not deal with changes from gm. to pounds.

Now in electricity we encounter four equations connecting six unknowns, m, μ, e, ϵ (this is the dielectric constant), i and A. These equations are

$$\text{force} = \frac{m_1 m_2}{\mu d^2}, \quad \text{force} = \frac{e_1 e_2}{\epsilon d^2}, \quad \text{force} = \frac{mids \sin\theta}{Ad^2} \quad \text{and} \quad i = \frac{de}{dt}.$$

It is usual for us to forget A because in the electrostatic and the electromagnetic systems A is unity, but A may not be unity in all systems.

The equation $c^2 = A^2/\mu\epsilon$ is not a fresh relation—as is often assumed. It is a derived statement coming from the above four equations, and should therefore be omitted.

Now, whenever we have four equations connecting six new quantities we must define new arbitrary units for two of the quantities. We always do this by defining these new units by reference to the values of a particular substance, i.e. vacuum. The unit of length is defined by reference to the distance between two scratch lines at the N.P.L. The unit of weight is defined by reference to the quality—mass of a particular substance, water—that is easily available. For the electrical quantities we choose for the particular arbitrary unit a more universal medium—a vacuum. In the e.s. and the e.m. systems $A = 1$ and either ϵ_0 or μ_0 is made unity also. Two out of the three must be fixed arbitrarily with arbitrary standards, so that *two* new units should be introduced into our theory of dimensions if it is to retain all its potential richness. We can, if we like, omit A and use only four. But if we do, we can solve problems involving four and not five unknowns. We will not be able to change our values to other systems in which A is other than unity. We have lost something by being lazy.

So it is with thermal units. We have one equation,

$$\text{heat} = SM\theta,$$

connecting three new quantities—heat, S and θ. We must therefore fix arbitrary units to two of these—generally heat and θ. When we do this we should introduce these two quantities in our dimensional equations. If we introduce only θ, our dimensional theory cannot change from one set of units of heat to another. We will be able to change our derived units from ° C. to ° F., but we will not be able to change from the calorie system to, say, the work scale.

This meeting has now convinced me of the incorrectness of my previous argument that since heat = (mass of water)(temperature rise in ° C.), we are justified in neglecting heat in our dimensional equations. The calorie is a new arbitrary unit, defined by reference to ° C. and to the weight of water. It differs fundamentally in no way from the unit of mass—a new arbitrary unit defined by reference to a volume of water. If we omit it we are deliberately throwing away something which can be of use to us.

Mr. G. H. West. I am quite sure in my own mind that Mr. Hansel's ideas are right and that his method of dealing with elementary dynamics by starting with *mass* as a quantity of matter is the proper way to start with boys. If you begin with W/g they do not seem to grasp the fundamental principles and they are confused from the start. I would proceed in this way :—

Mass of body is the quantity of matter in the body.

Unit mass is the quantity of matter in a piece of metal selected for the purpose (the standard pound).

Forces are measured by the accelerations they produce in the unit mass.

Unit force is one that produces unit acceleration in unit mass.

Equal masses are those which balance each other on a beam balance.

Density is mass of unit volume.

I would abolish the use of the term specific gravity in school, and leave it for technical purposes only, for example, specific gravity of milk and beer.

Author's reply :

Paper I—Ohm's Law. I cannot think of any complete circuit to which Ohm's law can be applied accurately in the form *current=E.M.F./resistance*, the resistance being accurately constant.

Ohm's law, that for a metallic conductor under fixed physical conditions the ratio of P.D. to current is constant, applies accurately to metallic conductors but seldom applies with the same degree of accuracy in other cases. It is the exception rather than the rule. Of course, the definition of resistance as the ratio of P.D. to current applies in all circumstances, whether Ohm's law is applicable or not, but, in general, this ratio is not constant.

In an experiment to illustrate Ohm's law I do not think that it is necessary to calibrate the ammeter and voltmeter by a method independent of Ohm's law in order to illustrate or to " prove " it (if it can be proved or verified by a single series of observations). A law represents universal experience, and consistency of deductions from it is the only way of establishing a law. I do not consider that the experiment described would be improved by calibrating the ammeter and voltmeter by any particular method or by substituting a tangent galvanometer for the ammeter and a potentiometer or capillary electrometer, etc., for the voltmeter. I think it is legitimate to assume (1) that the instruments used are reasonably accurate, (2) that a voltmeter measures P.D. and an ammeter measures current quite independently of any method used to graduate these instruments. Ammeters and voltmeters are graduated by methods which secure consistency of readings of different instruments calibrated by different methods, and the methods of calibration used to-day use the experience of a past generation in ensuring that the instrument shall read what it is intended to read. The same instruments may be used to show that Ohm's law applies to a metallic wire and that it does *not* apply to a radio valve. If the metallic wire did not conform to Ohm's law, the measurements would reveal the nature and extent of the departure from it. The experiment described in my paper *does* illustrate that for the particular wire under the conditions of the experiment the current is proportional to the P.D. applied to the wire, assuming that the voltmeter reads volts and the ammeter reads amperes.

I do not favour fantastic experiments specially designed to avoid instrument calibrations depending on Ohm's law, e.g. a separately-excited shunt dynamo with constant field running at variable speed, so that the P.D. at the brushes may be varied and measured with an electrostatic voltmeter and the load current with a Kelvin balance.

Paper II—Specific Heat. I do not agree with Prof. Stead's preference of the ratio definition of specific heat and his rejection of heat per unit mass per degree.

In the method of mixtures :—

Heat lost by hot substance = Heat gained by cold substance
= so many heat units (H) after substituting for s_0.

Hence, $s=H/m(t_2-t_1)$, i.e. heat units per unit mass per degree.

Similarly in the method of cooling.

The procedure I advocate is to discard the term *thermal capacity* as unnecessary, and to use the term *specific heat* for what Prof. Stead calls the thermal capacity per unit mass. Then :—

(1) Define unit quantity of heat. There are many ways of doing this (see Article *Calorimetry*, Encyc. Brit., 10th or later editions, by H. L. Callendar).

(2) Define specific heat (s) as the quantity of heat per unit mass of the substance per degree. With consistent units, the specific heat is the same in all units, e.g. in calories per gram per °C. as in B.Th.U. per pound per °F.

If the temperature of mass m of a substance of specific heat s changes from t_1° to t_2°, the heat change is $ms(t_2-t_1)$heat units.

(3) Define latent heat of fusion or vaporization (L), in a similar way, as heat required to change the state of unit mass of the substance at the fixed temperature of its M.P. or B.P. Fahrenheit latent heats are 9/5 times Centigrade latent heats.

Heat for a change of state $=mL$, at a fixed temperature.

(4) Change of entropy $(d\phi)=dH/T^\circ\text{K.}$, $dH=ms\,.\,dt$, and $\Delta H=mL$.

I do not agree that the method of mixtures or the method of cooling is only capable of interpretation on the assumption that s/s_0 is measured—this is one way of looking at it, but not the only way. On the ratio definition of specific heat, any heat equation involving both specific and latent heat would contain terms of different dimensions, e.g. the equation for finding the temperature when ice is added to a warm liquid.

The heat engineer and the student of thermodynamics are not particularly interested in the number of times the thermal capacity of a substance is greater than that of water. Nor do I think that students of physics, chemistry and engineering should regard quantity of heat as a " nebulous quantity ", nor that the equality between two quantities of heat should be concealed in the equations for the solution of heat problems.

As Prof. Stead points out, electrical methods give thermal capacity per unit mass per degree (which I call specific heat) which can be expressed in ergs (or joules) per unit mass per degree if J is known, or in heat units per unit mass per degree (using a double determination) if J is not known; and this is possible whichever definition of specific heat is adopted.

In my opinion, the type of definition which Prof. Stead recomends for rejection is to be preferred to the ratio definition for the following reasons :—

(1) Specific heat defined as heat per unit mass per degree enters into heat equations in a natural way, dimensionally correct, and without obscuring the nature of the thermal process.

(2) It is not essential to define unit quantity of heat as the thermal capacity of unit mass per degree of some standard substance. We may use the mean calorie between the ice and steam points (thus eliminating difficulties of thermometry and variation of specific heat); we may use as heat unit the quantity of heat required to melt unit mass of ice or some definite fraction of it, or the quantity of heat produced by a current of one ampere flowing through a resistance of one ohm, etc. Specific heat *not* defined as a ratio may be defined in terms of any unit of heat without any intermediary such as thermal capacity per unit mass per degree.

(3) If specific heat is a pure number, what is the expression for the change of entropy or internal energy of a substance when its temperature changes from $t_1°$ to $t_2°$? The ratio definition leads to an equation which is dimensionally unsound.

I do agree with Prof. Stead when he says " It seems to me a matter of little importance whether the term specific heat is used to express a ratio or a dimensional quantity; but it is unfortunate that it should be used indiscriminately in both senses." I do not agree that the two different names are needed, " as in the strictly parallel case of specific gravity and density ". The cases are not strictly parallel. For one thing, density varies in numerical value with the system of units used, whereas specific heat used synonymously with thermal capacity per unit mass is the same in all consistent systems of unit.

In this connection an article by Prof. Ferguson on " Some Units and Definitions " which appeared in the *S.S.R.*, March 1937, vol. xviii, no. 71, pp. 347–351, is relevant. Prof. Ferguson concludes his article as follows :—

" To sum up : it would seem, then, that there is a distinction between thermal capacitance, as defined above, and specific heat, defined as a ratio; that thermal capacity should be expressed in calories per gram per degree, or some equivalent unit, and that a name should be given to the unit of specific heat, if the ratio definition be retained—it is merely an accident of definition that the number which measures a specific heat is the same as the number which measures a thermal capacitance; or, alternatively, and preferably, that the term specific heat be retained for what we have here called thermal capacitance, and that the ratio-definition be abandoned ".

I find myself almost, but not quite, in agreement with the whole of what Prof. Ferguson says, and it seems relevant to this discussion to compare our points of view.

The equations of physics differ from those of pure mathematics in that the terms of a physical equation (i.e. the quantities separated by $+$, $-$, and $=$ signs) are dimensional—each term consisting of a dimensionless number and a unit having the same dimensions for each term. The equations of pure mathematics contain dimensionless numbers only.

Physical quantities such as length, time, mass, force, current, etc., are continuously variable magnitudes which may assume all values within the range of physical experience. The measurement of these magnitudes involves :—

(1) The selection of a suitable unit U of the same dimensions as the physical quantity to be measured.

(2) The determination of the number of times n that the physical quantity is greater than the unit.

n is a dimensionless number—it is the ratio of two quantities having the same dimensions. The measurement is then expressed in the form :—quantity $Q = nU$.

Experiment indicates that certain relations exist between physical quantities. For example :—

1. The quantity of heat H absorbed or rejected by a substance during a change of temperature is proportional to its mass m and to the change of temperature $(t_2 - t_1)$:

$$H \propto m(t_2 - t_1) \qquad \ldots\ldots (1)$$

or

$$H = km(t_2 - t_1). \qquad \ldots\ldots (1\,a)$$

2. The mass of a substance m is proportional to its volume v :

$$m \propto v \qquad \ldots\ldots (2)$$

or

$$m = kv. \qquad \ldots\ldots (2\,a)$$

3. The quantity of heat H produced by a current I flowing through a conductor of resistance R for t seconds :

$$H \propto I^2Rt \qquad \ldots\ldots (3)$$

or

$$H = kI^2Rt. \qquad \ldots\ldots (3\,a,$$

4. The pressure p in any specified unit is related to the depth x in any specified unit, and the density of a fluid d in any specified unit :

$$p = kxd. \qquad \ldots\ldots (4\,a)$$

The constant k in equations (1 *a*) and (2 *a*) is a physical constant for any one substance at a definite temperature and pressure. k is not a pure number but has dimensions.

Its value jumps discontinuously in passing from one substance to another. The value of such constants is not determined by comparison with a unit of the same kind, but by measuring the physical quantities involved in its defining equation, e.g. mass and volume in equation (2 *a*). My point of view here differs, I think, from that of Prof. Ferguson. I see no reason for using a unit of density or of specific heat. If units of mass, volume, temperature and heat have been defined, no further units of density or specific heat are necessary. Physical quantities are measured by comparison with a standard unit. Physical constants are evaluated by means of physical equations, which are also defining equations.

The constants k in equations (3 *a*) and (4 *a*) are constants of conversion and are dimensionless. The value of such constants is adjusted to suit the units of measurement of the physical quantities in the relation. The introduction of conversion constants into a relation to adapt it to fit chaotic units is a chaotic procedure, useless, confusing and unscientific.

Specific gravity (the ratio of two densities), specific heat (if defined as the ratio or two quantities of heat), coefficient of friction (the ratio of two forces), expansion ratio (the ratio of two volumes), efficiency (the ratio of two quantities of energy), etc., are dimensionless numbers or numerical coefficients, and need not appear in the equations of physics except where it is convenient to use a single symbol instead of two symbols and a division sign, e.g. expansion ratio, or where the ratio is used as a numerical coefficient such as efficiency or coefficient of friction.

The discussion here given relates to the physics of material systems in which quantities of matter, distances and time intervals are comparable in magnitude with those of ordinary human experience and perception. It does not follow that the same principles and treatment apply to the physics of minute space and time intervals (e.g. sub-atomic physics) to the ether (e.g. radiation, cosmic rays, etc.), or to the phenomena of light, electricity and magnetism (e.g. dielectrics and magnetic media). The discussion of these phenomena in terms of the ordinary theory of dimensions taking length, time, mass and temperature as indefinables, has led to much divergence of point of view. It may be that the same indefinables are now insufficient or irrelevant.

Paper III—Mass. Prof. Crowther has put his finger right on the spot in his reference to the insularity of some of the teaching in dynamics (see also Paper I, *Mass*, Appendix 4). Dynamics, taught as an isolated subject, leads to a confusion of ideas with regard to mass and inertia, or it may lead to the avoidance of the mass concept or even its complete elimination. To the physicist, chemist and engineer, the law of conservation of mass is of fundamental importance. This is not the law of conservation of inertia.

The mass of a body is the quantity of matter in it (Newton). Mass may be measured in any reasonable way whatever—weighing, titration, impact, inertial acceleration, gravitational acceleration, etc. The results by different methods agree. The terms *mass* and *inertia* are not synonymous. Inertia is only one of the properties of matter by which the measurement of mass is made possible, and this method is by no means the most convenient or the most accurate. I prefer the views of Newton to those of Mach or Poincaré (see Q. 74). Mach identifies mass and inertia, Newton does not.

The term *moment of inertia* is not well conceived. Nor are its equivalents, *second moment* or *second mass moment*. Why moment? Is this the changed meaning of *importanza*, the word used by Leonardo da Vinci? Moment of inertia measures rotational inertia. I think it would not only be unwise to pander to the vagaries of the popular usage of words and substitute inertia for mass, but the procedure would be inaccurate, unscientific and unsatisfactory.

What does a beam balance *measure*? It will be profitable to forget about rapidly-changing gravitational fields in space or time, or balances with beams of length to be reckoned in parsecs and of such great mass that the masses refuse to remain in the scale pans. Primarily, a beam balance compares moments: $m_1g_1d_1 = m_2g_2d_2$. If $d_1 = d_2$, it *compares* weights. If, in addition, $g_1 = g_2$, it *compares* masses. If the weighing is performed at the place where the standard weight is defined and with standard weights (which is never the case) it *measures* weight, but in no other case. If the weighing is performed with standard masses, as is always the case, it *measures* mass. The beam balance gives consistently the same results everywhere. How, then, can it *measure* weight, which is not the same everywhere? If it does not measure weight, how can it measure the tension in a string or any other force except with standard weights, where no-one ever weighs? So far as the beginner in science is concerned, a pair of scales has always measured mass (quantity of stuff) and not weight, but he has always called it weight and not mass. Is it not urgent to correct this misuse of the terms mass and weight? This is no more difficult than in the parallel case of heat and temperature.

The distinction between mass and weight should be taught early and at the very beginning of the science course. No slovenliness of speech or written record should be tolerated: there must be rigorous insistence on the correct use of the terms mass and weight. The teacher of mechanics should be careful to continue the work already begun in the science course.

" Mass in water " has been criticized. I entirely agree that this statement is thoroughly bad—mass is invariable. But is not the statement " weight in water " worse still? A beam-balance weighing is consistently the same everywhere. How can it measure weight or gravitational force, which are not the same everywhere? Apparent mass and apparent weight are even worse. Is the apparent weight of a body resting on a table nothing? What is the apparent weight of a picture hanging by an oblique cord? Surely the body has its natural weight independently of any particular method of support. In an Archimedes' experiment we are dealing with an unaltered weight and an additional force (the upthrust).

For my part, I do not think that this is a suitable subject for a beginner. At a later stage, when weight has been developed as a resultant of parallel forces acting at a centre of gravity, when composition and resolution of forces have been considered, and when the dependence of pressure on depth and density have been considered, the principle of Archimedes can be discussed and stated in the form that a body immersed in a fluid experiences an upthrust equal to the weight of the fluid displaced, and this upthrust acts at the centre of buoyancy. These ideas, coupled with the proportionality of mass and weight, will enable a student to state his problems intelligently and accurately, using the word mass or weight appropriately. Even a beginner cannot fail to notice the inconsistency of habitually using the word *mass* for beam-balance weighings and then suddenly using the term *weight* instead. This is also to be found in most text-books.

Density (or its reciprocal, specific volume) is the quantity that enters dimensionally into the fundamental equations of physics and engineering, e.g. the relation for the velocity of a sound wave through the air. Specific gravity is relative density, and a specific-gravity bottle may equally well be called a relative-density bottle if the volume of the bottle is not marked on it. But if the volume of the bottle is specified, this is for the purpose of determining density, and it is a density bottle (although I do not use this term myself):

$$\begin{aligned} \text{Specific gravity} &= \text{density of substance/density of standard}, \\ &= (m/v)/(m_0/v_0), \\ &= m/m_0 \text{ if } v=v_0, \qquad \dots\dots (1) \\ &= v_0/v \text{ if } m=m_0. \qquad \dots\dots (2) \end{aligned}$$

(1) is useful when dealing with a body that sinks;

(2) is useful when dealing with a body that floats.

The specific-gravity bottle, hydrostatic balance and Hare's apparatus determine relative density. If one of the liquids is the standard substance, the relative density becomes the specific gravity, but not otherwise.

If a beginner *must* deal with Archimedes' principle, then it should be presented experimentally. The difference between weighings of the body in air (really in a vacuum) and in a fluid is equal to the *mass* of fluid displaced. This is what an experiment with a beam balance discloses. That a beam balance measures mass and not weight is a matter of fact and not a matter of opinion.

Mr. Lauwerys complains of ambiguity and obscurity in the following statements :—

1. It is advisable to teach a student the conception of mass and not to ignore it completely.

2. That whatever form of treatment is followed, it will not be necessary to change over to some different type of treatment at a later stage. It would be better to defer the teaching of dynamics if the student is unable to grapple with its fundamentals.

3. Every student should be able to work in C.G.S. and in British units. All the better if he can work in any units—in fact, if he is properly trained he will be able to use all rational units with equal facility. Whatever point of view anyone may take on this matter, I think it will be conceded that if a type of training can be devised which enables all systems to be used with equal facility, or which applies equally to each of the rational systems of units, it is highly desirable that such a type of training be adopted universally. Any form of training which fits one system of units and does not apply to another, except by the introduction of unnecessary, unwanted and irrational units, should be discouraged.

There are four rational systems of units (see *Mass*, Appendix 5).

Rigorously and ruthlessly discard chaotic units such as the gram weight, grams weight per square centimetre. These irrational units are unnecessary, and in the same category as dyne-mass or poundal-mass.

A pressure may be expressed as a head of water or mercury, but avoid grams weight per square centimetre at all cost.

Taking Mr. Lauwerys' numbered points in order :—

1. The confusion of mass and inertia has been considered—the water-carts hardly merit it.

2. When Mr. Lauwerys says " we do not have to change over to a new system " and " we extend ", I refer to this as " a change over to some different type of treatment ". This may seem " obscure " to Mr. Lauwerys, but Section 6.5 of the Report of the Mathematical Association (p. 26) begins : " Those who commence with the use of gravitational units should, if they continue their studies sufficiently far, change over to the habitual use of absolute units (except for working certain types of numerical examples). It is recommended that the change should be effected in two stages, a preliminary stage in which absolute units are introduced and used for the first time, while gravitational units still remain dominant, and a final stage in which absolute units become habitual and gravitational units are merely reserved for problems to which they are specially suited."

Prof. Alfred Lodge's remarks are amusing when quoted against the use of $P=ma$, for he has most successfully revealed the kind of tangle that ensues when weight W is introduced into a simple problem in which it is irrelevant. Experiment shows that, for a given body, mass and weight are proportional, the constant ratio or the coefficient of acceleration being found to be equal to the mass of the body measured in other ways. Hence $P=ma$. Where is the relevance of W? Having established that $P=ma$ for forces in any direction, obviously $W=mg$.

Did Prof. Lodge state " The equality $P/W=a/g=$const. for a given body is more abstruse and should be delayed a bit " ? If so, I suggest it is so abstruse as to be incorrect, and should be delayed indefinitely.

I am not at all clear why Prof. Filon's remarks have been quoted. Does Mr. Lauwerys propose to *define* mass as W/g; or does he contend that it is not necessary to distinguish between mass and weight ? I follow Newton in defining mass as quantity of matter, and in measuring mass with a pair of scales. The mass (quantity of stuff) is the only real knowledge of things which is unavoidable in ordinary experience. That a pair of scales measures mass or quantity of stuff has been verified over and over again in the common incidents of daily life.

3. The irrational, unwanted units are the gram weight and grams weight per square centimetre, etc. $P/W=a/g$ cannot be applied to absolute units, and necessitates the introduction of gram weight, etc. How does Mr. Lauwerys propose to define unit pole and unit current ? How is the gram weight to be used in the quantitative study of electricity and magnetism ? I only use the gram weight when compelled to do so (lever experiments, calibration of spring, with beginners). I should very much prefer megadyne or kilodyne weights.

4. The difficulties of Karl Hahn (and of his pupils) are understandable if they approach the concept of mass through Attwood's machine. The difficulties are doubly distilled if kilograms and kilograms weight are used as the units of mass and force.

5. Mr. Lauwerys cannot fail to have convinced everyone of the futility of discussing mass with beginners according to the methods he proposes. The remedy is to follow Newton.

6 and 7 have been answered by Prof. Crowther, Prof. Stead and Messrs. Brereton, Dyson and West.

The formula $P/W=a/g$ was introduced by Prof. Greenhill and used in teaching applied mathematics to military cadets and officers. He completely ignored the conception of mass, and dealt exclusively with gravitational units and weight. The questions set in the Army Entrance Examination for many years dealt only with gravitational units. Public schools with a military side specialized in this type of training, and not a few of the text-books of mechanics have emanated from masters in these schools. Once a student has become habituated to this type of training it becomes impossible to change his point of view. He has difficulty with the conception of mass, and this difficulty arises entirely out of his previous training. Actually it is not mass that he finds difficult, but how to introduce this quantity into the equation $P/W=a/g$. Give me a young untrained beginner and I have no difficulty in teaching the conception of mass, but in the case of the habitual user of $P/W=a/g$ I find it impossible to teach him anything—his invariable reaction is a blind application of the formula $P/W=a/g$.

The remarks by Prof. Crowther and Prof. Stead on the force equation and units are well worthy of study. I include in this remark Prof. Crowther's remarks on the slug—his thrust at me is justified. I agree that the placation of the "practical" people has been a bad blunder, maybe irretrievable, maybe not. For this reason I have argued "to do a great right, do a little wrong". The introduction of the slug would enable those who desire to add to the number of units to go their own way without queering the pitch for absolute units.

Finally, I am extremely grateful to all those who have contributed to this discussion.

PRESENTATION OF DUDDELL MEDAL

THE eighteenth Duddell Medal was presented to Dr. W. D. Coolidge at a dinner of the American Physical Society held at Washington on 1 May 1942. The presentation was made by H.M. Ambassador, VISCOUNT HALIFAX, who said:—

It is a great pleasure for me to be here to-night, and for such a purpose as is the occasion of our meeting.

The Physical Society has decided to award its 18th Duddell Medal to another eminent American investigator, Dr. William David Coolidge, who has for many years past directed the research work of the General Electric Company. You must not think that in requesting me to make this award on its behalf the Physical Society was under any illusion that it was choosing a representative qualified to speak in its name about Dr. Coolidge's work. My knowledge of modern science, I am afraid, is shamefully slight, but I can truthfully claim that my interest in it is perhaps greater than that of many who profit by the results of science in daily life in such ways as using a car, listening to the radio, or getting about the country by air. Moreover, when the Society's invitation reached me I naturally took pains to consult with such learned minds as I could find, in order to inform myself to the best of my ability about the nature and significance of the achievements for which Dr. Coolidge is being honoured to-night.

As a result of my enquiries I am able to state that this award has been made primarily in recognition of his invention of the type of x-ray tube which is now in general use both in physical investigations and for medical purposes throughout the world, and which is everywhere known by Dr. Coolidge's name.

Dr. Coolidge's discoveries place him high in the list of holders of the Duddell Medal—a list which includes such honoured names as those of Sir Charles Vernon Boys, Sir Ambrose Fleming, Professor Ernest O. Lawrence and the late Professor Albert A. Michelson.

In constant stream the world has received incalculable benefits from the work of scientists, and the frequency with which those benefits have been conferred upon us has steadily increased over the past two centuries. Since Newton discovered the law of gravity the work of scientists in every land has combined to transform almost out of recognition the world in which, for so many thousand years, man had been living his brief, uncertain and arduous life. The field of medicine has been particularly favoured, and I am told that few, if any, living scientists could match the benefit which Dr. Coolidge has conferred upon medical science in making the x-ray tube the most important accessory of medical diagnosis. Diagnosis has indeed been revolutionized, and the number of lives saved by x-ray inspection is beyond calculation. It can also be said with gratitude that Dr. Coolidge's invention has reduced the hazards to which the earlier workers in the field of x rays were exposed, and to which, alas, so many of them sacrificed both health and life.

Although Dr. Coolidge already holds many honours, both American and British, this has seemed to the Physical Society an opportune moment to honour his exceptional achievements once more, since not only does the Coolidge tube in wartime mean the difference between life or death to thousands of wounded, but his researches in ductile tungsten have led to a complete revolution in signalling and communications, as radiotelephony has become indispensable to any form of modern warfare, and in the field of communications is the chief medium through which the Allies are bound together.

The Officers and Council of the Physical Society rightly feel that the award of the Duddell Medal to Dr. Coolidge has to-day a significance deeper than the honouring of the achievements of a great investigator. For the men of science of our two democracies are united by a common belief in liberty of thought and expression, without which their work would lack both opportunity and meaning.

This war, I need hardly remind so distinguished a company, is different from all others, not merely in the vastness of its geographical extent nor simply in the fundamental and all-encompassing nature of the principles for which we fight. That is all true—but there is more than that. It is a war in which the whole apparatus of Western civilization is called into play to save or to destroy the essence of that civilization itself. Many, I know, find cause for little short of despair when they contemplate that fact. Science, they say, has made all this and science is now destroying it. Where and how, they ask, are we ever to break that vicious circle ?

Yet I do not by any means despair of seeing that question answered, for we now begin to see that the old opposition between natural science and the world of the spirit was unreal. The material world about us has expanded into an universe which scientists themselves admit only faith can comprehend. In this new and infinite world that has been opened for us, scientist and layman alike face the unknowable again. Realizing as we do the narrow limits of our human range, we shall, I think, come once more to the knowledge that the issue of our endeavours, each in his own walk of life, lies in wiser than human hands. Thus, I believe, it is not unreasonable to hope that the old conflict will resolve itself upon a higher plane and mankind go forward again with a new humility and a re-awakened faith.

Dr. Coolidge, on behalf of the Physical Society it is my pleasure to present to you the Duddell Medal for 1942, and the certificate of award and the honorarium that go with it.

In accepting the Medal, Dr. Coolidge said* :—

May I first thank Your Excellency for your great courtesy in coming here and giving me the pleasure of receiving this honour at your distinguished hands. This is generous compensation for the disappointment which I naturally felt because present conditions made it inadvisable for me to attend a meeting of the Physical Society in London to receive the award directly from their President.

I am deeply conscious and appreciative of the honour of being made a Duddell Medallist, and for several reasons. The medal takes high lustre from the great scientific achievements of Duddell, in whose honour the Physical Society established the memorial fund—a lustre enhanced by the eminence of that society itself and by the conspicuous merit of former recipients. It carries special significance and value for me in coming from England, the home of my ancestors, the source of many great achievements in science, and the stalwart guardian of what are to me the finer things of life.

To return for a moment to the work of Duddell. Had he done nothing more than develop the magnetic oscillograph, his contribution to the electrical science and art would have been outstanding. One cannot visit a laboratory today where electrical researches or electrical engineering developments are in progress without being greatly impressed by the importance of the rôle played by the oscillograph, which yields so much essential information obtainable in no other way. For frequencies within its range, the magnetic oscillograph, substantially as developed by Duddell, is still in general use. For higher frequencies, the cathode-ray oscillograph is used, but here also credit must be given to Duddell, for his magnetic oscillograph served to blaze the trail. The art of radio also is heavily in debt to Duddell for his discovery of the singing arc, his design of a high-frequency alternator, and of his thermo-ammeter, all of which contributed to the early development of wireless telegraphy.

To make clear the limitations of my part in the work on which the award was based, I must briefly outline the conditions under which the work was done.

Our laboratory, at the time the ductile-tungsten work was started, was only five years old. This laboratory sprang from the vision of Dr. E. W. Rice and his associates. He was then in charge of General Electric engineering and had seen clearly the extent to which the products of our factories had been derived from the results of scientific research, such as had been carried out notably by English physicists.

The laboratory was started in the hope that it might contribute new facts and principles on which new products could be based. It was one of the earliest industrial research laboratories in this country and was established primarily for fundamental research. It was itself an experiment, but an experiment carried out under extremely favourable conditions, as both its sponsors and its director, Dr. W. R. Whitney, were men of great courage, optimism, and faith in the power of scientific research methods. The strongest single psychological factor conducing to the success of the laboratory was the spirit of co-operation which Dr. Whitney was able to build up not only within the laboratory but also between the staff and the engineers of the Company.

---

* Reproduced by courtesy of the American Institute of Physics from *J. Appl. Phys.* **13**, 478 1942).

That co-operation was so close and so generous and so effective that a very large share of the credit for the ductile-tungsten and x-ray work belongs to those who were associated with me in that work. Credit is also due to the courageous and liberal management of the Company in giving its continued support to such work as the ductile-tungsten effort, which over a long period of years was exceedingly expensive and of questionable promise.

While the factory process for making ductile tungsten is to-day fundamentally the same as that developed in the laboratory, it has undergone constant improvement and been given increased utility by our engineers for the various applications which have been found for the product.

Although this process did not constitute the first application of powder metallurgy, it seems to have been at the time the most important one which had been made, and for this reason it may perhaps be fairly credited with having been in no small measure responsible for the subsequent development of the art of powder metallurgy now widely applied to other metals, and to mixtures of metals and refractory compounds.

My participation in the work on x rays, which forms the other basis for the award, followed as a result of the keen interest which I had felt in x rays ever since the announcement of their discovery, also from the work which we did in adapting wrought tungsten to use as a target in the earlier x-ray tube in place of platinum, and from Dr. Irving Langmuir's studies on the thermionic emission of electrons from tungsten filaments in high vacuum. I had the advantage also of intimate first-hand knowledge of the limitations of the x-ray tube of that time and of the source of those limitations. Stable and controllable thermionic electron emission rendered unnecessary the gas content which alone had been responsible for the most troublesome limitations of the earlier tube. At the same time it provided the much desired flexibility of control and operation. I was also fortunate in having readily available not only tungsten wire for the cathode filament but also large masses of wrought tungsten which were at the time indispensable for the target. For when our first successful tubes were made, we did not know that it would ever be possible sufficiently to de-gas a composite target of copper and tungsten such as had been generally used in the earlier type of tube.

I cannot forego this opportunity to pay my tribute to Sir Owen Richardson who, more than any other, had laid the groundwork for our knowledge of thermionic emission.

From the beginning, it was clear that the high-voltage generating equipment which had been developed for use with the earlier tube was ill-suited to the hot-cathode type, and for this reason later developments of the hot-cathode tube and of the high-voltage source have, even up to the present time, gone hand in hand.

While some hot-cathode tubes of the first so-called universal type are still in use, this design has generally been supplanted by special tubes giving greatly improved performance for various specific uses.

For example, types have been developed for operation on unrectified alternating current. The characteristics of the tube have permitted various forms of equipment in which the tube and the entire high-voltage circuit are completely enclosed within grounded metal, thus eliminating all danger of electric shock and facilitating the attainment of adequate x-ray protection. For radiography, definition has been greatly improved in various ways, notably by the use of elongated focal spots and, more recently, by this combined with rapid rotation of the target.

For medical therapy and for industrial radiography, tubes for higher voltages than could be used with the first design have been developed, first by merely increasing dimensions. Then, because of troublesome field currents arising from too high potential gradients at the cathode, we were led to depart from the simple two-electrode structure and to use a multiplicity of tubular accelerating electrodes in addition to cathode and anode. This apparently makes it possible to provide tubes for voltages as high as can be produced.

A recent development in the generation of high voltage to supply such tubes is the resonance transformer. This has no iron within the primary and secondary coils, thus permitting the x-ray tube to be located here, where it is electrostatically shielded by the surrounding coils. The size and weight of such equipment have been greatly reduced by substituting for oil a compressed gas having high dielectric strength.

For x-ray diffraction work, the recent development of tubes with linear focus combined with metallic beryllium windows results on the one hand in great reduction of the required exposure time, and on the other in a considerable extension in the long-wave-length direction of the available spectrum.

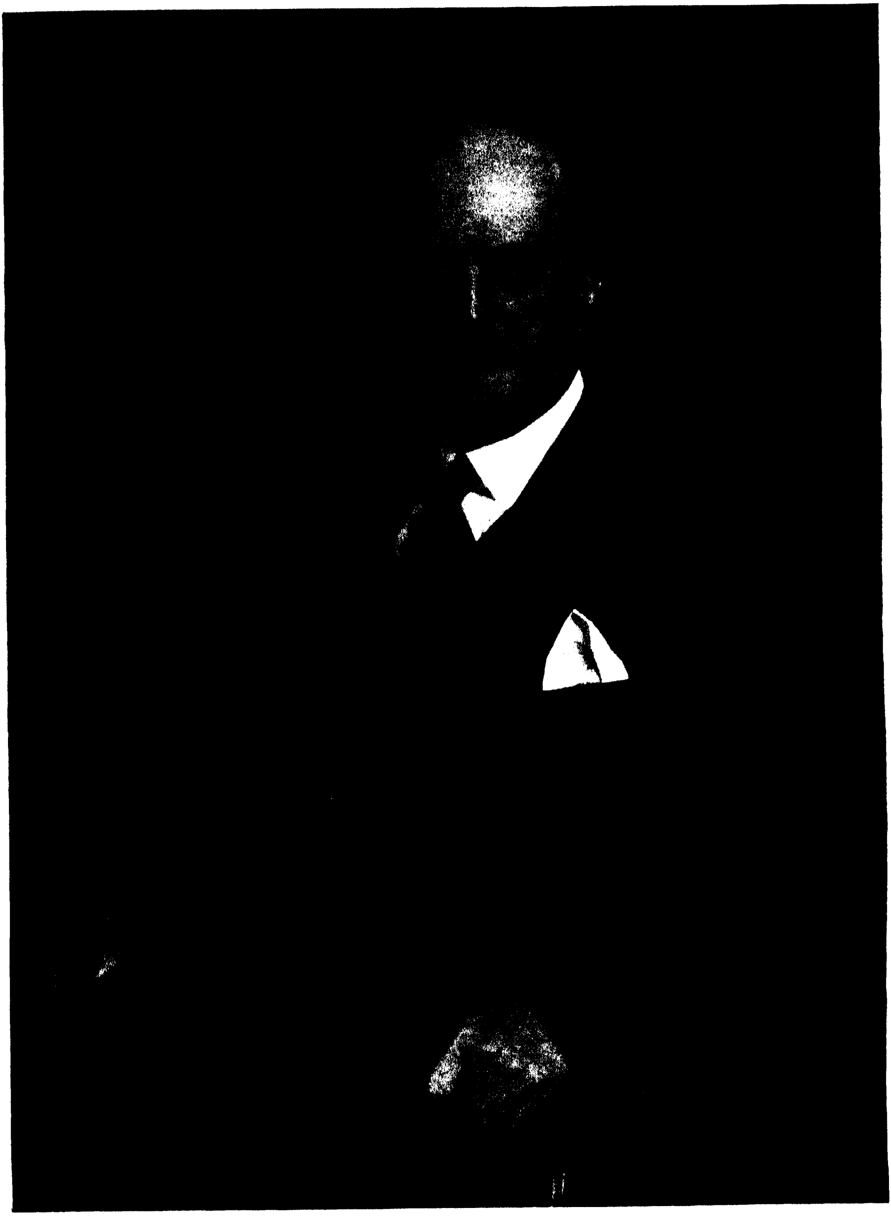
It is exceedingly pleasant to think of any useful purpose subserved by war. World War No. 1 did greatly accelerate both the development and the introduction to use of the hot-cathode tube. The present war has already led to the extensive use of the photo-roentgenographic method for the chest examination of recruits for the armed forces, and this, because of its greatly reduced cost, excellent results, and convenience, is certain to find widespread use after the emergency. The present war is also stimulating enormously the application of x rays to industrial uses, both radiographic and fluoroscopic. As an illustration of this, although the first portable million-volt radiographic x-ray outfit went into use only about a year ago, the fortieth one will soon have been completed. These million-volt outfits make possible the rapid examination of even as difficult subjects as steel castings and welds having a thickness of as much as eight inches, and thereby are contributing most helpfully to many industrial war activities. As another example, fluoroscopic outfits are being built and used for the rapid examination of small aluminium die castings, thus permitting immediate rejection of many whose defects could otherwise be seen only after a considerable amount of machine work had been performed on them.

To me, Your Excellency, the greatest value of this award lies in the fact that it is one of the many evidences of the cordial and friendly spirit which has for years existed between British and American scientists. This has been evidenced anew by the close and effective co-operation in the present war-work between the scientists of the British Commonwealth and those of this country. It has been greatly facilitated through your various war agencies in England, the National Research Council of Canada, and our Office of Scientific Research and Development. Your scientists are bringing us full reports of the vast amount of splendid work which you are doing in support of the war effort. We are making good use of it and are in turn keeping you informed concerning our work. This augurs well, both for the emergency and for the more leisurely pursuits to follow. This mutual aid in our common cause is strengthening the bonds between our nations, and we may hope that, when victory is won, those bonds may serve to hold us together in the task of rehabilitating the world on a firm basis of liberty, security and opportunity for all.

In every country of the world there are men of good will, who desire peace and friendly relations with other nations. Misrepresentative governments may gain temporary ascendancy, and by ruthless aggression and acts of cruelty bring upon themselves the hatred of their victims, but there remains, though temporarily submerged and silenced, the peace-loving mass of human beings who comprise the bulk of the nation, and who will respond in kind to an approach in the spirit of friendship and mutual helpfulness.

It is a far cry to-day, but we must certainly look hopefully to the time, and labour diligently for it, when such relations of mutual esteem and friendship as now exist not only between the scientists but also between the peoples of the British Commonwealth of Nations and the United States of America can spread to all peoples and to all countries.

In closing, Mr. President, I wish to express to the American Physical Society my gratitude for graciously providing this very pleasant setting for this occasion and thus lending endorsement to the honour conferred upon me by the Physical Society of London.



[Photo: J. Russell & Sons.

SIR WILLIAM BRAGG, 1862–1942.

[*Photo*: *Special Press*.

Sir William Bragg, with some of the audience for his Christmas Lectures on "The Universe of Light", at the Royal Institution, 1931–2.

OBITUARY NOTICES

SIR WILLIAM BRAGG

William Henry Bragg was born on 2 July 1862. His father, originally a master mariner, had in his later years become a farmer. Bragg was brought up on the farm and went to the village school. Later he was adopted by an uncle who lived at Market Harborough; then he went to King William College, Isle of Man. While there he won a scholarship at Trinity College, Cambridge, and this success determined his future career. He was Third Wrangler in the Mathematical Tripos in 1884. Having taken his degree, he had to look about for a position to earn as quickly as possible, and this circumstance may perhaps have prevented his entering for a fellowship. However this may be, his future career was, as he told me, determined by a curious accident. Being on his way to some academic function he had forgotten his gown. He hurried back to fetch it, and on his second journey he encountered J. J. Thomson in the street, who happened to mention that the post of Professor of Mathematics and Physics at Adelaide University was vacant owing to the resignation of the late Prof. Horace Lamb. The decision was imminent and Bragg had only time to apply for the post by telegram. He was chosen, though he knew little or nothing of experimental physics at this time. He took out books on the subject and read them on the voyage.

He remained at Adelaide for twenty-two years, and married in 1889 the daughter of Sir Charles Todd, the Postmaster-General and Astronomer Royal of South Australia.

Bragg's career as an original experimentalist began very late in life. Indeed it would be difficult, if not impossible, to mention anyone else who achieved great success in this direction without having made an earlier start than he did. He was, in fact, forty-two years of age when his first paper appeared in the *Philosophical Magazine*.

The way in which his original experimental work arose was this. He had been asked to give a public lecture in illustration of the newly discovered wonders of radium, and in order to illustrate it, and perhaps for his own instruction, he had procured a sample of the rather weak radium preparations which were then on the market. Arranging some lecture illustrations with this, he was led to ask himself certain questions which started his career as an original experimentalist.

His fundamental idea was this. The intensity of ionization produced by the alpha rays of radium falls off with distance owing to the rapid absorption of these rays in the air. Bragg, using a thin ionization chamber, determined the rate of this fall off, and using suitable methods confined the rays to a sensibly parallel beam. The remarkable fact then emerged that there were a number of discontinuities in the curve representing the ionization in a thin layer as a function of the distance of that layer from the source. These discontinuities were traced to the presence of the various component radioactive disintegration products, which Rutherford had detected as normally present in a radium preparation. Each of these products had its characteristic range: that is to say, there was for each of them a characteristic distance to which the alpha radiation which it gave off could penetrate. There was this rapid falling off of ionization, the range being, in fact, a definite numerical constant, which could be specified for each of these component radioactive bodies. At the time when Bragg's early work began, such researches presented technical difficulties which it is not easy to envisage now that more powerful radioactive sources have become available. These early researches were made in collaboration with R. Kleeman, his demonstrator.

The value of Bragg's work was quickly recognized, and to some of the Cavendish Laboratory School it came as a surprise that a new star should have arisen in this way at the Antipodes.

Bragg left Australia in 1909 to take up the appointment of Professor of Physics at the University of Leeds, vacated by Stroud, of range-finder fame. The family returned home by the ill-fated ship *Waratah*. Her behaviour during this voyage, which was her first, gave cause for grave anxiety, and Bragg was often consulted by the commander during the voyage. The ship disappeared without a trace on her next voyage between Durban and Cape Town, and Bragg was able to give scientific evidence at the enquiry which was based on his observations of her stability, and was helpful in elucidating the probable causes of the ship's mysterious disappearance.

He only remained a short time at Leeds, being invited to take up the post of Professor at University College, London, in 1915. It was during this particular period in England that his most famous researches were made. He was quick to take advantage of the discovery by Laue of the diffraction of x rays by crystals. Laue's original point of view had been that the crystal acted as a three-dimensional grating, its spacing being such as to be suitable for the short wave-lengths of x rays. Bragg's earlier contributions were chiefly in the direction of applying this discovery to the examination of x-ray spectra. In this work he was associated with his son, and he was always careful to explain that the guiding idea of these researches, which assimilated each layer of atoms in the crystal to a reflecting surface, was due primarily to his son. These layers were uniformly spaced, and if the angle of reflection was varied, the reflected rays from the successive planes might or might not be in phase, according to circumstances. Owing to the large number of parallel reflecting surfaces which came into the question, the angle at which reinforcement was obtained was very critical and the resolving power correspondingly great. It was accordingly possible to determine the wave-length of x rays or, if this was regarded as given, to determine the spacing of the crystal planes with accuracy, and the subject suddenly jumped from a condition of vagueness and uncertainty to one of full precision with measurements making everything clear and definite within 1 % or less. He built the x-ray spectrometer, and one of its first successes in his hands was to prove the existence of x-ray spectra. He worked out the $h\nu$ relationship, and the critical absorption and its relation to atomic weight. In the meantime his son was working out crystal structures, and he himself turned more in this direction after the war. This field of discovery is too well known and descriptions too generally accessible to make it desirable to devote much space in this notice to elaborating the details. Bragg and his son have given in their books admirable summaries of their work and the development which has been built upon it by their successors. It was found possible to map out precisely the position and absolute distances of the atoms of the various elements in all the ordinary inorganic and organic crystals. The general conclusions of crystallography which had been gained by examining the angles of the crystal faces and classifying the various degrees of symmetry were all confirmed and illuminated by precise allocation of the positions of atoms in the crystal lattice. It would perhaps be fair to say that the bare bones of geometrical crystallography were clothed in flesh. Bragg and his collaborators went far to exhaust this field of research, and the work stands for all time as a strong and firmly built structure complete in itself.

During the war of 1914–18, Bragg was appointed a member of the Admiralty Board of Invention and Research under Lord Fisher, and applied himself wholeheartedly to the problem of the acoustical location of submarines. He was in charge of the experimental station, first at Hawkscraig and then at Harwich. In spite of the energy and patriotic enthusiasm expended in this work, the ultimate practical outcome was not great, and it is no secret that quite different methods are now employed.

In 1923 the position of Director of the Royal Institution became vacant by the death of Sir James Dewar, and Bragg was offered and accepted the post. It was commonly believed in scientific circles that after many years' experience he was tired of elementary teaching and the routine duties of an ordinary professorship, and was glad to escape into a different and in some respects less strenuous atmosphere.

The position of Director of the Royal Institution carried with it the tenancy of the flat at the top of the building in Albemarle Street, which in former times had been occupied by Faraday, by Tyndall and by Dewar, and here he made his home. The atmosphere of the Royal Institution was congenial to him. It was, as he said, a friendly place, and he was by natural gifts well qualified to carry out the social duties attaching to it, in the way of receiving the members after the Friday Evening Lectures, occasionally entertaining foreign scientific guests of distinction at dinner and otherwise, and generally contributing to maintaining the social traditions of the place.

It fell to him to give one or more courses of afternoon lectures in the course of the year. He was particularly successful in giving the famous Christmas Lectures, which had been initiated by Faraday and have remained an important feature of the programme of the Institution ever since. The courses he gave in this way were: "The World of Sound", "Concerning the Nature of Things", "Old Trades and New Knowledge", "The Universe of Light". These were afterwards published in book form.

It will be noticed that in selecting the topics for these and other courses he went considerably outside the range of his own speciality. Indeed, he rather leaned towards the older physics, which lent itself more easily to experimental illustration, with which he took great pains. He did not aim at spectacular successes of the music-hall type, or, to express it otherwise, his illustrations were directed to making the subject clear rather than to gaining personal kudos for the lecturer. He would carefully discuss beforehand with his entourage what degree of detail would be suitable to the audience and how far it would be safe to go beyond the easier and more superficial aspects of the subject. In particular he leaned to the use of models in making his topics clear. He was well acquainted with all the traditional methods in experimental illustration developed by Tyndall and others, and had evidently studied them in detail during the period of his professorship in Adelaide. He looked on everything from the point of view of the audience, and was careful never to be too far over the heads of the majority.

The care and trouble taken in collecting and arranging the material for these lectures was very great. Bragg was not content to rely on what could be found in books or in original printed sources. He also gathered information in factories or workshops. He estimated the time spent in preparing for one hour's lecture at not less than a month.

He took a great interest in the history of the Royal Institution. He gave lectures from time to time on a variety of topics in this connection. These included courses or single lectures on "Acoustical Problems treated by [the late] Lord Rayleigh", "From Faraday's Notebooks", "Two Old Friends of the Royal Institution (Spottiswoode and de la Rue)", "Benjamin Thompson, Count Rumford", "History of the Vacuum Flask", "Tyndall's Experiments on Magne-Crystallic Action", "Experiments from the Researches of Sir James Dewar". He was particularly interested in all Faraday relics and Faraday traditions, and gave much attention to the editing of Faraday's notebooks in connection with the centenary celebration in 1931, which was largely organized by him. Apart from the subjects already mentioned, his Royal Institution lectures dealt for the most part with topics connected directly or indirectly with x-ray crystal analysis.

In the course of Bragg's administration it was found necessary to make important structural changes in the Royal Institution. Some of those to whom the place was holy ground regretted the necessity for this, which seemed something of a desecration, but there is no doubt that Bragg acted with a due sense of responsibility and with the full backing of the officers of the Institution. In particular the rising seats of the theatre were originally carried on a structure of wood, and the managers of the Institution were advised that the dangers of fire raised possibilities which were terrible to contemplate, particularly in the case of an audience of young people as at Christmas time. They felt it impossible to carry the responsibility of this risk, and although the Local Authorities had no power to compel action, they strongly advised it. The theatre was reconstructed with iron supports for the seats. After this reconstruction it was in the main replaced as nearly as possible in its original form, the historic lecture table being retained and the shape of the room left unaltered. The old Physical Laboratories used by Tyndall and Rayleigh ceased to exist as such, and the place was remodelled to provide a new series of rooms, including improved lecture-preparation room, librarian's room and other conveniences. At the same time the Laboratories of the Institution were merged more completely with the Davy-Faraday Laboratory, and the machinery used by Dewar for the liquefaction of gases, which had now become obsolete, was removed, with the exception of a few small pieces of historical interest.

On the more serious scientific side arrangements were made by which a certain number of salaried workers of reasonable maturity could be available to work under Bragg in the Davy-Faraday Laboratory. The original design of Dr. Ludwig Mond, the founder of the Laboratory, had been that independent workers with means of their own would be able to come there to work on their own problems, but in practice the scheme hardly worked out in the way expected. There were not enough workers of the kind contemplated to make proper use of the facilities of the Laboratory, and it was very desirable that younger men who had not an independent income or position, but whose creative activities were still on the east side of their meridian, should be encouraged to work at the Laboratory, and this change was made possible by Bragg's reorganization. He collected the necessary funds from a variety of sources. Some large industrial concerns were generous contributors. Among the workers with him from University College were A. Müller and G. Shearer. Other well-known workers were W. T. Astbury, J. D. Bernal, C. F. Elam (Mrs. Tipper), R. E. Gibbs, K. Lonsdale, and several others of comparable distinction. They worked for the most part on the problems of crystal structures which Bragg and his son had initiated, developing these more particularly on the side of organic chemistry, and later with reference to the structure of animal tissues and the like.

The researches of Bragg and his school were made for the most part without specially powerful appliances. The high-power x-ray installations (5 kw. and 50 kw.) developed by Dr. A. Müller for the more exacting problems of crystal analysis were perhaps the only exceptions.

Bragg was President of this Society for the years 1920–22 and delivered an address on the structure of organic crystals, which he had recently investigated by x rays. From 1935 to 1940 he was President of the Royal Society. He felt, as he told me at the time, a sinking of the heart at having undertaken this latter office, apparently from an entirely uncalled-for feeling of inadequacy and doubt as to his ability to carry the burden at the somewhat advanced age of 73. In fact, however, his energies lasted well. He had a strong hold on the respect and affection of the Fellows, and this was by no means the last of the public duties he was called on to perform. Towards the end he had to husband his strength, and usually took a taxi even for the short journey from Albemarle Street to Burlington House.

He was created K.B.E. in 1920 and given the Order of Merit in 1931, and was awarded at various times the Rumford and Copley Medals of the Royal Society. He died at the Royal Institution on 12 March 1942. RAYLEIGH.

WILLIAM ALFRED BENTON

BORN 16 March 1871, eldest son of William Benjamin Benton of Handsworth, William Alfred Benton was educated at King Edward's School, Birmingham. He received his engineering training with Frank Stacey & Co., milling engineers, of Birmingham, and then practised as a consulting engineer in Birmingham. About this time he became interested in automatic weighing machines and took out a number of patents.

During the last war he erected and worked a by-products plant for the extraction of ammonia and important organic compounds from waste material.

After the war, in 1921, he joined W. & T. Avery, scale-makers, of Soho Foundry, as head of their Research Department.

He had a great knowledge of the theory and practice of weighing, and in recent years put together for W. & T. Avery a unique collection of photographs, models and documents relating to the history of weighing from the earliest times. A few years ago he gave a demonstration to the Society on this subject. Amongst many other devices he invented an instrument for the rapid determination of specific gravities.

Benton was interested in optics and had patented an epidiascope, as well as carrying out some experimental work on photographic lenses. He was a keen student of literature and history, and some years ago, under the pen-name of John Gentleshaw, published a long poem, *The Observing Angel*, dealing with the economic depression.

He had been a fellow of the Physical Society since 1924, and was also a member of the Chemical Society, the Royal Photographic Society, the Newcomen Society, and the Birmingham Metallurgical Society, of which he was President in 1933–34. This Society had also awarded him their Silver Medal.

He died suddenly at his house in Handsworth on 19 June 1942.

FRANCIS JAMES SELBY

F. J. SELBY, an early Fellow and Member of Council of the Optical Society, and also for long a Fellow of the Physical Society, died on 5 March. By training a mathematician, he was first a student of University College, London, and later of Trinity College, Cambridge. Some of his recollections of that time will be of interest, at least to a number of the older members of the Society. He was very decided in his view that Karl Pearson was the best mathematical lecturer he had ever met. At Cambridge he graduated as sixth wrangler in 1891, a year after Miss Fawcett was placed above the Senior Wrangler. It happened that Selby knew somewhat intimately the work of Miss Fawcett and of some men in the same year at Cambridge, for they had worked together at University College. However able Miss Fawcett may have been, she had been easily surpassed by one of the men who went from University College to Cambridge. At Cambridge, while the one worked steadily, the other occupied himself in other ways, and took a high place in the tripos on the work he had done before going up. Selby was sure that had he given a moderate amount of time to mathematics at Cambridge he would have been far ahead of everyone else in his year.

After leaving Cambridge, Selby became mathematical master first at Bristol Grammar School and later at Repton. In 1903, two years after its inauguration, he joined the staff of the National Physical Laboratory to take charge of the prediction of tides in the Indian Ocean, then about to be transferred to the Laboratory, and to start an Optical Division intended to extend the range of optical work beyond that carried out in the Observatory Department at Kew. To these duties Selby added that of Secretary to the Director, an office which absorbed his time increasingly. The excellent relations he established with the whole of the Laboratory staff, combined with his good judgement and administrative gifts, rapidly led him to a key position in the Laboratory organization. This was only formally recognized in 1918, when he was appointed Secretary of the Laboratory, but in fact the general business of the Laboratory had been largely in his hands for several years: this greatly assisted the Laboratory in giving such notable service to the country throughout the war period 1914–19, when the Director, Sir Richard Glazebrook, was called on to spend much of his time elsewhere.

In 1909 the Advisory Committee for Aeronautics was formed, with Glazebrook as Chairman and Selby as Secretary. These additional duties led Selby to relinquish charge of the Optics Division. He continued as Secretary of the Committee until, in 1919, it was replaced by the Aeronautical Research Committee. He was awarded the C.B.E. in recognition of his services. He retired from the Laboratory in 1932 on reaching the age of sixty-five.

Selby had set himself the task of writing the early history of the National Physical Laboratory on his retirement. His knowledge of both men and events marked him out as the ideal writer of such an account. Unfortunately, before the work was completed he had a stroke which affected his memory, and the work was unavoidably brought to an end.

Throughout his service at the Laboratory Selby maintained his interest in mathematics, giving particular attention to relativity and the quantum theory. He took charge of the scheme instituted at the Laboratory for the continued education of junior members of the staff and for securing the best industrial experience for mechanics trained in the Laboratory workshops. In all his work he was inspired by a high sense of the service that the Laboratory should render to the Nation, and set a great example of unselfish devotion, often in circumstances of peculiar difficulty. T. S.

BERNARD MYNOTT NEVILLE

NEVILLE was born in 1886, and may be said to have spent his whole life in the atmosphere of the school. From 1899 to 1905 he was a pupil at William Ellis School, and he joined its staff in the latter year. Meanwhile, he continued his own studies, and after winning a scholarship at London University, he took his degree in 1909. He joined the Physical Society in 1910, and often attended its meetings. His own interest, however, was in the more academic bodies ; he was treasurer of the S.M.A., and then advertisement manager of the *School Science Review*, and served long on its Advisory Committee. For 22 years he had been an examiner in Physics in the University of London.

That he lived such a full life, and was always cheerful, despite the handicap of lameness from which he had suffered since the age of three, is no small tribute to his spirit and character.

He died on 11 April 1942, leaving a widow, whom he had married as recently as December 1940.

REVIEWS OF BOOKS

Mass Spectra and Isotopes, by F. W. ASTON. Second Edition. Pp. xii + 276. (London: Edward Arnold and Co., 1942.) 22*s*. 6*d*.

The appearance of another edition of a book which occupies as historic a place as does this one in our academic and scientific literature is a most welcome event. The five chapters comprising Part I (" Historical ") of this edition and the first four chapters of Part II (" Production and analysis of mass spectra "), while having been brought up to date, are not radically different from the first edition, but a tenth chapter, on " Modern high-power mass spectrographs ", has been added, and includes those developments in experimental technique which have occurred during the important decade which has just elapsed. The title of Part III, " The elements and their isotopes ", remains unchanged, but its form is completely altered. Instead of following the groups of the periodic system, it is now arranged according to atomic number, and a wealth of information, based on results published up to 1941, is contained therein. Part IV (" Theoretical and general ") is much the same in form, but has been extensively revised. This second edition is a most worthy and useful successor to the first.

W. B. M.

Time and the Universe: a new Basis for Cosmology, by [the late] F. L. ARNOT. Pp. 76, with plate. (Sydney: Australasian Publishing Company, Ltd., 1941.)

In 1937 the late Dr. F. L. Arnot, then a lecturer in Natural Philosophy in the University of St. Andrews, discussed with me a theory of time and cosmology of which a short account was published in *Nature* on 25 June 1938. This he subsequently expanded into a monograph which was ready for publication before he left Scotland in 1939 for his home university in Sydney. Here he rewrote and rearranged large sections in the endeavour to simplify and improve the presentation of his thesis. After his death in October 1940, publication was made possible through the generosity of the Scientific Publications Committee of the Commonwealth of Australia, and Prof. O. U. Vonwiller has attempted to carry out as editor what he believed to be Dr. Arnot's wishes as regards alterations.

According to Prof. P. W. Bridgman, the essential point to consider in defining any physical magnitude is the *process* of measurement—the rules of operation by which we obtain the measuring numbers associated with the physical phenomenon. When we apply this method to time, perhaps the most fundamental of all human conceptions, it is reasonable to inquire into the possibility of more than one time scale. For instance, there may be one scale for atomic phenomena, and another for ordinary mechanical phenomena. Such an assumption was made by de Sitter and appears in the cosmological theories of both Milne and Dirac. The atomic time-scale may be called the t-scale, while the scale obtained by using an ordinary macroscopic time keeper may be called the tau (τ) scale. This τ-time is measured by our ordinary clocks regulated by a pendulum or by a rotating planet. We then choose another time, called t-time, which is connected with τ-time by a logarithmic relation. It is then shown (not assumed) from the resulting theory that an atomic clock keeps this particular t-time, and that the age of the universe is finite. In *The Nature of the Physical World*, Eddington has said : " The difficulty of an infinite past is appalling. It is inconceivable that we are the heirs of an infinite time of preparation ; it is not less inconceivable that there was a moment with no moment preceding it ". To this Arnot replies : " Since neither time scale possesses any intrinsic claim to preference, we have two equally valid pictures of the universe. In one picture the universe was created at a definite epoch ; in the other the universe has existed for an infinite time. Both views are equally correct."

The title of Chapter I is " The number of particles, mass and radius of the universe ", that of Chapter II, " The bridge between cosmological theory and quantum theory ". In some respects the next chapter, on the red-shift of nebular spectral lines, is the most important in the book. The interpretation of this as a Doppler effect due to recession of the nebulae is a speculation which has unsatisfactory features. The present theory explains the red-shift as due to a decrease in the velocity of light in t-time, and the formula obtained for the degree of red-shift as a function of the distance of the nebula is in excellent agreement with the experimental observations. Moreover, a more decisive test of the theoretical formula will be possible when results have been obtained with the new 200-inch reflector.

Chapter IV is mainly mathematical, being concerned with relativistic t-mechanics, and the concluding chapter reviews certain selected subjects, such as the number of particles in the universe, the gravitational constant, and finally the relation between the two time-scales.

In the history of science, instances abound of the early statement of a new principle or theory being received with scant attention or even with active opposition. Time will show whether this new presentation of its own nature is of value in removing outstanding difficulties and in suggesting new investigations. H. S. ALLEN.

Spectroscopy and Combustion Theory, by A. G. GAYDON. With a foreword by A. C. EGERTON, F.R.S. Pp. x + 191. (London: Chapman and Hall, Ltd., 1942.) 17*s*. 6*d*. net.

As a result of the systematization of molecular spectra on the basis of the quantum theory during the first decade or more after the last war, many of the band systems observed in the spectra of various flames were definitely attributed to emitting radicals, such as C_2, CN, CH, NH and OH, whose existence had hitherto been almost, if not entirely, unknown. At the same time the actual states of electronic and vibrational excitation of such emitters in flames and other sources of radiation were completely determined and described. Subsequent investigations led to an understanding of the presence and the rôles of these intermediate products in combustion processes and have yielded much new and valuable information on flame reactions.

The results of these recent applications of theoretical and practical spectroscopy to combustion problems are presented and discussed in this excellent monograph by the author of a short report on the same subject in the recent volume (vol. **8**, 1941) of our Society's *Reports on Progress in Physics*. The book is the first of its kind and can be unreservedly recommended to all who are engaged upon or interested in such problems.

An elementary introduction to the theory of molecular spectra is given in the opening chapter, and the conditions under which band systems appear in emission are discussed in the second chapter to provide an adequate theoretical background for the investigations, both qualitative and quantitative, described in later chapters. In the next six chapters the spectra of the hydrogen flame, hydrocarbon flames, cool flames, atomic flames, the carbon monoxide flame, explosions of hydrocarbons as in the internal combustion engine, and inorganic flames are described in some detail. Ch. IX deals with absorption spectra as a means of studying the progress of chemical reactions, as in slow combustion of hydrocarbons. Accounts of flame spectra in the infra-red region (Ch. X) and the calculation of life-times of activated molecules (Ch. XI) form a basis for discussions of the author's own work on the after-burning of carbon monoxide (Ch. XII), and of measured and calculated flame temperatures (Ch. XIII). Ch. XIV is devoted to the derivation of dissociation energies, and the final chapter to the application of spectroscopic results to the kinetics of the combustion process.

The usefulness of the monograph is enhanced by an Appendix of 20 pages giving numerical data and other information on molecular spectra and energy levels occurring in flames, by two plates of excellently reproduced flame spectra, and by an adequate index. W. J.

College Physics, by WILLIAM T. MCNIFF. Third Edition. Pp. 657. (New York: Fordham University Press, 1942.) $4.00.

The two volumes of the previous edition of this book are now combined into one very attractive volume with extensive revisions and much new material. Like several recent American books, it gives a little of almost everything in the vast field of physical knowledge, from elementary mechanics of solids and fluids to modern nuclear physics. The text is broadly divided into seven parts: Mechanics (159 pages), Heat (68 pages), Sound (19 pages), Light (107 pages), Magnetism (17 pages), Electricity (198 pages), Modern Physics (38 pages). There follow a number of Appendices containing simple mathematical notes, formulae and tables, and information concerning the chemical elements and the electromagnetic spectrum. A large selection of numerical exercises, solved and unsolved, is included.

One of the author's main objects is to lay a foundation for more advanced study in each branch of the subject. Modern ideas are introduced wherever possible, not only in Part 7, which is entirely devoted to investigations of the last half-century, but in other parts of the text ; indeed, the subject matter of three whole chapters of Part 6 (Ch. 45 on conduction of gases, x rays, thermionic emission, rectifier tubes, the cathode-ray oscillograph, isotopes and the mass spectrograph ; Ch. 46 on radioactivity and cosmic rays ; Ch. 47 on radio telegraphy and telephony, television and the electron miscrocope) might well have been transferred to Part 7, which at present consists of one chapter (Ch. 49, on atomic structure, quantum theory and nuclear physics). These four chapters form perhaps one of the best features of the book. Two others which call for special mention are Chapters 43 and 48, which give elementary accounts of generators and motors and alternating currents.

In the Introduction, the author stresses and exemplifies the dependence of medical science on the applications of physical principles, and in two chapters (rightly described in the preface as foreign to the usual text-book of physics) he outlines some such applications, namely, physiological and therapeutical effects of changes in air density (Ch. 5), and electricity as a therapeutic agent (Ch. 44). These two, and several other chapters, render the book especially suitable for the medical student.

A few minor slips have been noticed : the ray-tracing in fig. 99 (D) on p. 277 is wrong; in problem 28 on p. 468 " kilowatt " should be " kilowatt-hour " ; on p. 510 " e/n " appears in one place for " e/m ". It seems unfortunate that while the protons-and-neutrons structure of atomic nuclei is represented in the very useful and informative chart on pp. 607–8, the earlier conception, protons-and-electrons, is mentioned alone on p. 580 and as an alternative on p. 610.

W. J.

Wave Guides, by H. R. L. Lamont. Pp. vii + 102. (London : Methuen and Co., Ltd., 1942). 4*s*.

The study of wave guides has recently become one of considerable importance because of their many practical applications with very short electromagnetic waves. During the last few years many authors have written accounts of work, both theoretical and experimental, on this subject, unfortunately not all using the same nomenclature. The present volume deals with the essential aspects of the theory, and includes most of the important work on the subject to date, with a uniform set of units and nomenclature.

The monograph is about 100 pages in length, with six chapters. The first chapter, after a brief historical survey, deals with the fundamental electromagnetic theory, and then with the rectangular guide, for which type the analysis is most simple. In the next chapter the author discusses general transmission theory, circular guides, dielectric guides, guides of elliptic and other special types of cross-section : the horn is also considered. Chapter III is devoted to the study of the attenuation and stability of waves in the guides; the losses due to the finite conductivity of the walls and to the conductivity of the dielectrie are calculated. In chapter IV the author deals with the question of the wave impedance of a guide; then, on the basis of the simple laws of geometrical optics, a physical picture of the propagation of waves along a guide is given. The last two sections are devoted to the use of wave guides as resonators and as radiators. Resonators of various geometrical forms are compared and the relevant energy relations developed. The applicability of guides and horns to the production of beam radiation is indicated in the last chapter.

As the author mentions in his preface, the experimental side has received scant mention; an excellent classified bibliography is provided, however, so that reference to the experimental technique is facilitated. Dr. Lamont is to be congratulated on producing such a useful summary of the more important elements in wave-guide theory.

J. A. S.

The Scientific Journal of the Royal College of Science. Vol. XII (Session 1941–42). (London: Joint Publication Board of the R.C.S., Imperial College Union, 1942.) 5*s.* 6*d.* cloth, 4*s.* 0*d.* paper.

It is the first duty of a reviewer to give his readers some notion of the content, scope and execution of a book. The content of such a work as this can be judged rightly only by listing the topics treated, and here, as briefly as may be, is a statement of the published lectures, given in the Session 1941–42 before the three Societies under whose aegis the Journal is published:—

CHEMISTRY: Electron Microscopy (A. L. G. Rees). Methane as an Engine Fuel (A. C. G. Egerton). Arrhenius, Ostwald and van't Hoff (A. Findlay). The Chemistry of Hashish (A. R. Todd). The Acetylene Alcohols (A. W. Johnson).

NATURAL HISTORY: Luminescence in Decapod Crustacea (R. Dennell). The Nature of Viruses (R. K. S. Wood).

MATHEMATICS AND PHYSICS: Graphical Methods in Teaching and Research (A. Ferguson). Electron Diffraction and Surface Structure (B. W. Soole). Some Physical Aspects of the Chemical Bond (A. Charlesby). The Differential Analyser (R. E. Beard). Diffraction Methods and Molecular Structure Determinations (A. Charlesby).

An admirably varied programme indeed, with fare to suit very different tastes. As regards scope and execution, the lectures are designed, and rightly designed, to appeal to the student trained in the subject of any particular lecture; they give him brief and clear surveys of the recent advances in that subject. But they do more than this. They give to the chemist an opportunity of learning something of advances that have been made in other branches of sciences, and conversely, workers in these other branches may learn something of the doings in the world of chemistry. This is a most useful function of the Journal, and one for which the present reviewer is happy to have this opportunity of placing on record his gratitude.

The volume is well produced, is, like its predecessors, thoroughly interesting; and there is nothing cheap about it, save its very modest price. A. F.

The Calculation and Design of Electrical Apparatus, by W. WILSON, D.Sc., M.I.E.E. Third Edition. Pp. xv + 240. (London: Chapman and Hall, Ltd., 1941.) 10*s.* 6*d.* net.

The second edition of this book, dated 1940, has already been reviewed in these pages. The speedy issue of a third edition evidences the serviceableness of the book, which has been amplified here and there, especially in the chapters on Heating and Heavy Conductors. D. O.

RECENT REPORTS AND CATALOGUES

Chance-Parsons Optical Glass. (Catalogue CP 400.) Pp. 19. CHANCE BROTHERS AND Co., LTD., Glass Works, Smethwick, near Birmingham.

Studies on Adhesives: Part I, *Ground-nut protein adhesives for plywood*, by D. NARAYANAMURTI, V. RANGANATHAN and P. K. BASU ROY CHAUDHURI; Part II, *Ground-nut protein—formaldehyde dispersions as plywood adhesives*, by D. NARAYANAMURTI and K. SINGH. (Indian Forest Leaflets, nos. 15 and 16.) Pp. 9 and 13. FOREST RESEARCH INSTITUTE, Dehra Dun, U.P., India.

Protection of Radium during Air Raids. (National Bureau of Standards Handbook H 38, May 1942.) U.S. DEPARTMENT OF COMMERCE, Washington, D.C. 10 cents.

Optical and Mechanical Characteristics of 16-millimeter Motion-Picture Projectors, by R. E. STEPHENS. (National Bureau of Standards Circular C 437, June 1942.) Pp. 22, U.S. DEPARTMENT OF COMMERCE, Washington, D.C. 10 cents.

INDEX

PAGE

Absolute measurement of electrical resistance by a method using the average electromotive force of a commutating generator 27
Achromatic reflection microscope for use with visible or ultra-violet light, (Demonstration of photomicrographs.) 288
Aluminium, thermal expansion of: further experiments 487
Approximate solution for the distribution of temperature or potential with cylindrical isothermal or equipotential surfaces, An 63
Atomic disintegration experiments, High-voltage apparatus for 439

Balls, E. G., *see* Nettleton, H. R.
Barrier dimensions, influence of, on the vibrational spectrum of ethane 471
Benham, W. E.: The nature of temperature 121, 289
Benjamin, Mark (Obituary notice) 297
Benton, William Alfred (Obituary notice) 535
Beta-ray spectrometry, The application of thick magnetic lenses to 266
Blackman, M.: Note on Raman's theory of the specific heat of solids 377
Blood-pressure measurement, The physics of 211
Born, Max: Lattice dynamics and x-ray scattering 362
Bragg, Sir William H.: The secondary x-ray spectrum of sylvine: Theory and experiment 354
Bragg, Sir William H. (Obituary notice) 531

Canal-ray positive-ion sources 245
Charlesby, A.: Influence of barrier dimensions on the vibrational spectrum of ethane 471
Charlesby, A.: Note on the comparison of the subsidiary maxima in electron diffraction and x-ray diffraction patterns from single crystals 379
Cinema screens, The whiteness of 98
Colour blindness and its importance in relation to industry 219
Colour mixing in the dyeing industry 1
Colour physics, Research on, at South Kensington, 1877–1942 303

Coolidge, William David: Eighteenth Duddell Medallist 209, 527
Coulson, C. A.: Diamagnetic susceptibility of methane 51
Craggs, J. D.: A high-voltage apparatus for atomic disintegration experiments 439
Craggs, J. D.: Canal-ray positive-ion sources 245
Crystal dynamics, x-ray study of: An historical and critical survey of experiment and theory 314
Cylindrical isothermal or equipotential surfaces, An approximate solution for the distribution of temperature or potential with 63

Diamagnetic susceptibility of methane 51
Dielectric loss in paraffin-wax solutions 422
Differential photo-electric photometer for precision work, A rotating 66
Dimensions, electric and magnetic 46, 289
Dimensions of physical quantities 428, 504
Distribution of temperature or potential with cylindrical isothermal or equipotential surfaces, An approximate solution for the 63
Duddell Medallist, Eighteenth 209, 527
Duncanson, W. E.: Note on some recent papers on physical quantities and their dimensions 504
Dyeing industry, Problems of colour mixing in the 1

Eddington, Sir Arthur: The theoretical values of the physical constants 491
Electric and magnetic dimensions 46, 289, 428, 504

PAGE

Electrical resistance, Absolute measurement of, by a method using the average electromotive force of a commutating generator 27
Electrometer for measurement of voltage on small ionization chambers . . 435
Electron beams, Space-charge effects in velocity-modulated 135
Electron diffraction and x-ray diffraction patterns from single crystals, Comparison of the subsidiary maxima in 379
Electrostatic component of the force of sliding friction 14
Equipotential surfaces. An approximate solution for the distribution of potential with cylindrical 63
Ethane, Influence of barrier dimensions on the vibrational spectrum of 471
Evans, David S. and Mendelssohn, K. : The physics of blood-pressure measurement 211
Evans, R. C. and Peiser, H. S. : A machine for the computation of structure factors 457
Expansion of aluminium, Further experiments on the thermal 487

Farmer, F. T. : An electrometer for measurement of voltage on small ionization chambers 435
Flow, A problem in two-dimensional 128
Friction, Electrostatic component of the force of sliding 14
Fröhlich, H. : Dielectric loss in paraffin-wax solutions 422
Fuchs, W. H. J. and Kompfner, R. : On space-charge effects in velocity-modulated electron beams 135
Fundamental laws and definitions in physics : I—Ohm's Law ; II—Specific heat and Newton's law of cooling ; III—Mass ; Discussion . 151, 159, 164, 509

Garfitt, D. E. M. : The measurement of impulse voltages by means of small sphere gaps 109
Grundy, J. G. : Problems of colour mixing in the dyeing industry . . . 1

Halifax, Viscount : Presentation of the eighteenth Duddell Medal . . . 527
Hampton, W. M. : Some problems relating to optical glass 391
Hansel, C. W. : Fundamental laws and definitions in physics : I—Ohm's law ; II—Specific heat, etc. ; III—Mass; Discussion . . . 151, 159, 164, 509
Harris, Tom (Obituary notice) 298
Harrison, V. G. W. : The measurement of near-whites in the paper industry . 86
Heys Hallett, C. G. : The whiteness of cinema screens 98
High-voltage apparatus for atomic disintegration experiments 439
Holmes, J. G. : The nature and measurement of whiteness. 81
Hopper, V. D. : Thermostats employing external surface control. 55

Impulse voltages, The measurement of, by means of small sphere gaps. . . . 109
Influence of barrier dimensions on the vibrational spectrum of ethane . . . 471
Ionization chambers, Electrometer for measurement of voltage on small . . 435
Isothermal surfaces, An approximate solution for the distribution of temperature with cylindrical 63

Johnson, B. K.: Photomicrographs illustrating the paper "An achromatic reflection microscope for use with visible or ultra-violet light", **53**, 714 (1941). (Demonstration.) 288

Kompfner, R., *see* Fuchs, W. H. J.

Lattice dynamics and x-ray scattering 362
Lonsdale, Kathleen : X-ray study of crystal dynamics : An historical and critical survey of experiment and theory 314

PAGE

Macey, H. H. : A problem in two-dimensional flow 128
MacGregor-Morris, J. T. and Stainsby, A. G. : A rotating differential photo-electric photometer for precision work. (Discussion.) 66
Machine for the computation of structure factors 457
Magnetic lenses, The determination of their optical properties, and their application to beta-ray spectrometry 266
Mayneord, W. V. : The measurement of radiation for medical purposes . . . 405
Mendelssohn, K., *see* Evans, David D.
Methane, Diamagnetic susceptibility of 51
Microscope, An achromatic reflection, for use with visible or ultra-violet light (Demonstration of photomicrographs.) 288

Near-whites, The measurement of, in the paper industry 86
Nettleton, H. R. and Balls, E. G. : The absolute measurement of electrical resistance by a method using the average electromotive force of a commutating generator 27
Neville, Bernard Mynott (Obituary notice) 537

Obituary notices 297, 531
Optical glass, Some problems relating to 391
Optical properties of thick magnetic lenses, The determination of the . . . 266

Paper industry, The measurement of near-whites in the 86
Paraffin-wax solutions, Dielectric loss in 422
Peiser, H. S., *see* Evans, R. C.
Photometer for precision work, A rotating differential photo-electric 66
Photomicrographs illustrating the paper "An achromatic reflection microscope for use with visible or ultra-violet light". (Demonstration.) 288
Physical constants, Theoretical values of the 491
Physical quantities, dimensions of 46, 289, 428, 504
Pitt, F. H. G. : Colour blindness and its importance in relation to industry . . 219
Positive-ion sources, Canal-ray 245

Radiation, The measurement of, for medical purposes 405
Raman's theory of the specific heat of solids, Note on 377
Ratio of density of town gas to that of air at the same temperature and pressure, The smoke-manometer determination of the 278
Recent reports and catalogues 80, 302, 387, 541
Resistance, Absolute measurement of, by a method using the average electromotive force of a commutating generator 27
Reviews of books (*see* Index, p. 546) 70, 205, 300, 388, 463, 538
Rheinberg, Julius (Obituary notice) 297

Schnurmann, Robert and Warlow-Davies, Eric : The electrostatic component of the force of sliding friction 14
Searle, G. F. C. : The determination, by smoke manometer, of the ratio of density of town gas to that of air at the same temperature and pressure 278
Secondary x-ray spectrum of sylvine : Theory and experiment 354
Selby, Francis James (Obituary notice) 536
Selwyn, E. W. H. : The dimensions of physical quantities 428
Siday, R. E. : The determination of the optical properties of thick magnetic lenses, and the application of these lenses to beta-ray spectrometry 266
Single crystals, Comparison of the subsidiary maxima in electron diffraction and x-ray diffraction patterns from 379
Sliding friction, Electrostatic component of the force of 14
Smoke-manometer determination of the ratio of density of town gas to that of air at the same temperature and pressure 278

PAGE
Solids, Raman's theory of the specific heat of 377
South Kensington, Research on colour physics at, 1877–1942 303
Space-charge effects in velocity-modulated electron beams 135
Specific heat of solids, Raman's theory of the 377
Spectrum of ethane, Influence of barrier dimensions on the vibrational . . 471
Stainsby, A. G., *see* MacGregor-Morris, J. T.
Structure factors, Machine for the computation of 457
Subsidiary maxima in electron diffraction and x-ray diffraction patterns from single crystals, Comparison of the 379
Susceptibility of methane, Diamagnetic 51
Sylvine, Secondary x-ray spectrum of : Theory and experiment 354

Temperature, The nature of 121, 289
Theoretical values of the physical constants 491
Thermal expansion of aluminium : Further experiments 487
Thermostats employing external surface control 55
Two-dimensional flow, A problem in 128

Velocity-modulated electron beams, Space-charge effects in 135
Vibrational spectrum of ethane, Influence of barrier dimensions on the . . 471
Voltage on small ionization chambers, Electrometer for measurement of . . 435
Voltages, The measurement of impulse, by means of small sphere gaps . . 109

Warlow-Davies, Eric, *see* Schnurmann, Robert.
Whitehead, S. : An approximate solution for the distribution of temperature or potential with cylindrical isothermal or equipotential surfaces 63
Whiteness of cinema screens 98
Whiteness, The nature and measurement of 81
Wilson, A. J. C. : The thermal expansion of aluminium : Further experiments . 487
Wright, W. D. : Research on colour physics at South Kensington, 1877–1942 . 303

X-ray diffraction and electron diffraction patterns from single crystals, Comparison of the subsidiary maxima in 379
X-ray scattering, Lattice dynamics and 362
X-ray, Secondary, spectrum of sylvine : Theory and experiment. 354
X-ray study of crystal dynamics : An historical and critical survey of experiment and theory 314

Yarnold, G. D. : Note on electric and magnetic dimensions 46, 289

INDEX TO REVIEWS OF BOOKS

PAGE

Arnot, [the late] F. L.: *Time and the Universe: a new Basis for Cosmology* 538
Aston, F. W.: *Mass Spectra and Isotopes* 538

Bagenal, Hope: *Practical Acoustics and Planning against Noise* 390
Barrer, Richard M.: *Diffusion in and through Solids* 74
Bridgman, P. W.: *The Nature of Thermodynamics* 463
Buros, Oscar Krisen (editor): *The Second Yearbook of Research and Statistical Methodology Books and Revie* 70

Champion, F C *niversity Physics.* Part III, *Light* 300
Champio : *University Physics.* Part IV, *Wave Motion and Sound* 79
Chapm *Edmond Halley as Physical Geographer, and the Story of his Charts* 76
Compto hur H.: *The Human Meaning of Science* 470
Comrie, J.: *The Twin Marchant Calculating Machine and its Application to Survey Prob s* 300

Dampier, Sir William Cecil: *A History of Science* 467

Eyring, Henry, *see* Glasstone, Samuel.

Findlay, Alexan r: *Introduction to Physical Chemistry.* Second edition 469
Findlay, Alexan r: *Practical Physical Chemistry.* Seventh edition 73

Gamow, George: *Biography of the Earth: Its Past, Present and Future* 469
Gamow, George: *The Birth and Death of the Sun* 74
Gaydon, A. G.: *Spectroscopy and Combustion Theory* 539
Glasstone, Samuel, Laidler, Keith J. and Eyring, Henry: *The Theory of Rate Processes* 388

Hatfield, H. Stafford: *European Science* 205
Healey, R. H. and Reed, J. W.: *The Behaviour of Slow Electrons in Gases* 206
Holmyard, E. J.: *Endeavour.* Vol. 1, no. 1 207
Humby, S. R. and James F.: *Science and Education (Current Problems,* 15) 466

Kaye, G. W. C. and Laby, : *Tables of Physical and Chemical Constants.* Ninth edition 78

Laidler, Keith J., *see* Glasst , Samuel.
Lamont, H. R. L.: *Wave Guides* 540
Loeb, Leonard B.: *A Laborat y Manual of Electricity and Magnetism* 205

McNiff, William T.: *College Phy s.* Third edition 539
Mann, W. B. (editor): *Reports on Progress in Physics,* Vol. VIII 302
Middleton, W. E. Knowles: *Meteorological Instruments* 301
Middleton, W. E. Knowles: *Visibility in Meteorology.* Second edition 72

Noakes, G. R.: *A Text-Book of Electricity and Magnetism* 78

Parr, G.: *The Cathode Ray Tube and its Application* 208
Petterssen, Sverre: *Weather Analysis and Forecasting. A Textbook on Synoptic Meteorology* 75

Reed, J. W., *see* Healey, R. H.
Royal College of Science: *The Scientific urnal of the Royal College of Science,* Vol. XII 541

Stratton, J. A.: *Electro-Magnetic Theory* 77

Wenner, Ralph R.: *Thermochemical Calc* 207
Wilson, W.: *The Calculation and Design* tion 541
Wilson, W. Ker: *Practical Solution of Tor* cond edition 71

PRINTED BY TAYLOR AND FRANCIS, LTD.

I.A.R.I. 75

INDIAN AGRICULTURAL RESEARCH
INSTITUTE LIBRARY, NEW DELHI.

Date of Issue	Date of Issue	Date of Issue

GIPNLK—H-40 I.A.R.I.—29-4- 5—15,000